ON THE FLUCTUATIONS OF MATRIX ELEMENTS OF THE QUANTUM CAT MAP

LIOR ROSENZWEIG

ABSTRACT. We study the fluctuations of the diagonal matrix elements of the quantum cat map about their limit. We show that after suitable normalization, the fifth centered moment for the Hecke basis vanishes in the semiclassical limit, confirming in part a conjecture of Kurlberg and Rudnick.

We also study sums of matrix elements lying in short windows. For observables with zero mean, the first moment of these sums is zero , and the variance was determined by the author with Kurlberg and Rudnick. We show that if the window is sufficiently small in terms of Planck's constant, the third moment vanishes if we normalize so that the variance is of order one.

1. INTRODUCTION

The study of quantum wave functions of classically chaotic systems has been extensively studied in recent years. One well known result is that in the mean square sense the matrix elements of smooth observables concentrate around the classical average of the observable in the semiclassical limit [29],[32],[3]. This is known as the "Quantum Ergodicity Theorem". The problem of whether all matrix elements converge to the classical average (the "Quantum Unique Ergodicity" problem) has no general result so far. This has been extensively studied, and in some arithmetic cases both positive (when considering desymmetrized eigenfunctions) answers (cf [24, 11, 30, 19, 15]) and negative answers (cf [7, 15, 16]) have been given.

Another important property is the distribution of the matrix elements. It was suggested by Feingold and Peres [8] that for generic systems with D degrees of freedom, the variance of the matrix elements about their mean decays with Planck's constant $\hbar$ as $\hbar^D$, with a prefactor given in terms of the autocorrelation function of the classical observable. Furthermore in [6] Eckhart *et al* predict that after normalizing the fluctuations of the matrix elements, they have a limiting Gaussian distribution about their limit with the same expected value and variance. Some arithmetical models were found to deviate from these predictions [25, 22, 15].

In this paper we study properties of these fluctuations for the quantum cat map. To describe these properties we first recall the model.

1.1. **The Quantum cat map.** The quantized cat map is a model quantum system with chaotic classical analogue, first investigated by Hannay and

Berry [10] and studied extensively since, see e.g. [14, 4, 19, 7, 28]. While
the classical system displays generic chaotic properties, the quantum sys-
tem behaves non-generically in several aspects, such as the statistics of the
eigenphases, and the value distribution of the eigenfunctions [21].

We review some of the details of the system in a form suitable for our
purposes, see e.g. [4, 19, 28]. Let A be a linear hyperbolic toral automor-
phism, that is, $A \in SL_2(\mathbb{Z})$ is an integer unimodular matrix with distinct
real eigenvalues. We assume $A \equiv I \mod 2$. Iterating the action of A on
the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ gives a dynamical system, which is highly chaotic.
The quantum mechanical system includes an integer $N \geq 1$, the inverse
Planck constant, (which we will take to be prime), an N-dimensional state
space $\mathcal{H}_N \simeq L^2(\mathbb{Z}/N\mathbb{Z})$, and a unitary map $U = U_N(A)$ of $\mathcal{H}_N$, which is the
quantization A. The eigenvalues and the dimension of the eigensapces of U
are related to the order of A modulo N. Let $\mathrm{ord}(A, N)$ be the least integer
$r \geq 1$ for which $A^r \equiv I \mod N$. When N is prime the distinct eigenphases
θ_j are evenly spaced (with at most one exception) with spacing $1/\mathrm{ord}(A, N)$,
and in fact, the distinct eigenphases are all of the form $j/\mathrm{ord}(A, N)$. The
eigenspaces all have the same dimension (again with at most one exception)
which is $(N \pm 1)/\mathrm{ord}(A, N)$.

For fixed small $\epsilon > 0$, as $N \to \infty$ through a sequence of values such
that $\mathrm{ord}(A, N) > N^\epsilon$ all the matrix elements converge to the phase space
average $\int_\mathbb{T} f(x)dx$ of the observable f [20, 2] (However, note that there are
"scars" found for values of N where $\mathrm{ord}(A, N)$ is logarithmic in N, see
[7].) The condition on $\mathrm{ord}(A, N)$ is valid for most values of N (in fact
$\mathrm{ord}(A, N) > N^{1/2+o(1)}$ for almost all N, c.f. [20, Lemma 15]), Moreover, it
was shown by Kurlberg in [17] that assuming GRH, for almost all primes N
$\mathrm{ord}(A, N) \gg N/b(N)$ for any function $b(x)$ tending to infinity more slowly
than $\log x$, and for almost all values of N, $\mathrm{ord}(A, N) \geq N^{1-\varepsilon}$.

In [19] Kurlberg and Rudnick introduced a group of unitary operators,
the Hecke group, that commutes with U. It is shown in [19] that if $\{\psi_N\}$
is a sequence of Hecke eigenfunction (a joint eigenfunctions of U and all
elements of the Hecke group), then for any smooth function $f \in C^\infty(\mathbb{T}^2)$
the matrix elements $\langle \mathrm{Op}_N(f)\psi_N, \psi_N \rangle$ converge to the space average $\int_{\mathbb{T}^2} f$.
In [22] they raise a conjecture about the fluctuation of the matrix elements
around the limit for a fixed function. The operator $\mathrm{Op}_N(f)$ is decomposed
by the Fourier decomposition of f, that is if $f(x) = \sum_{n \in \mathbb{Z}^2} \hat{f}(n)e(nx)$, then
$\mathrm{Op}_N(f) = \sum_{n \in \mathbb{Z}^2} \hat{f}(n) \mathrm{Op}_N(e(nx))$. They conjecture that for fixed $0 \neq n \in \mathbb{Z}^2$ the set $\langle \mathrm{Op}_N(e(nx))\psi_j, \psi_j \rangle$ becomes equidistributed with respect to the
Sato-Tate measure, and after considering symmetries of the system these sets
become independent for different choices of n (a more precise explanation is
given in section 8). Agreement with this conjecture is shown in figures 1,2.
In figure 1 the cumulative distribution function (cdf) of the fluctuations
of the matrix elements for the fixed function $f(x) = e(x + y)$ is shown
compared with the cdf of a random variable with Sato-Tate distribution (the

probability density function in this case is $p(x) = \frac{1}{2\pi}\sqrt{4 - x^2}$). In figure 2 the fixed function is $f(x) = e(x + y) + e(x + 2y)$. In this case the expected limiting distribution is of the sum of two independent random variables with Sato-Tate distribution, and again the cdf of the matrix elements shows high agreement with the conjecture. The matrix used in both cases is $\begin{pmatrix} 7 & -2 \\ 4 & -1 \end{pmatrix}$
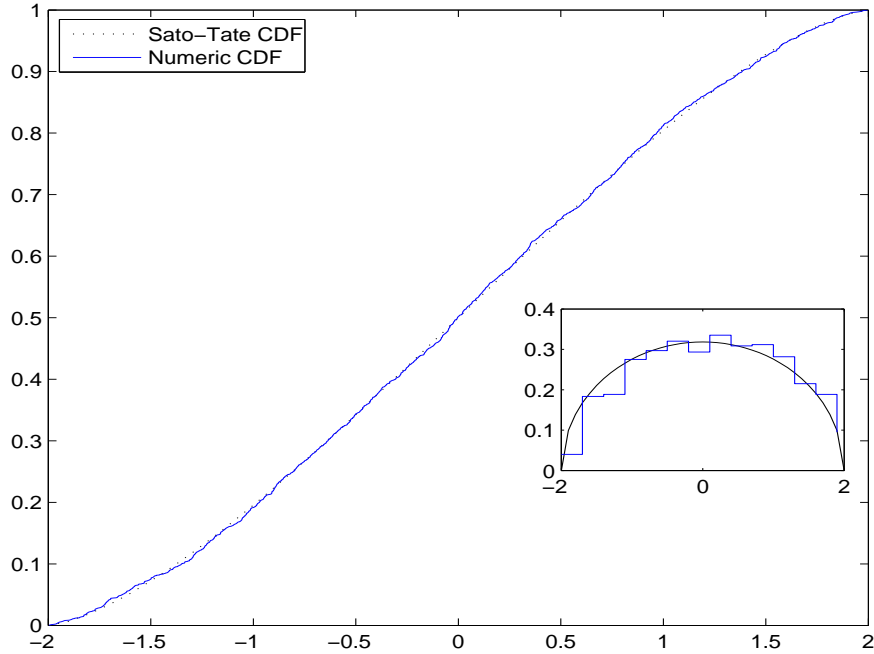

FIGURE 1. Cumulative distribution function of Hecke eigen-basis, with $f(x) = e(x+y)$, $N = 1997$ compared to Sato-Tate comulative distribution function

Another way to study the fluctuations of the matrix elements, is by studying the sum of diagonal matrix elements of $\mathrm{Op}_N(f)$ over eigenphases lying in a random window of length $1/L$ around θ. More generally we consider a window function, constructed by taking a fixed non-negative and even function $h \in L^2([-\frac{1}{2}, \frac{1}{2}])$ and setting $h_L(\theta) := \sum_{m \in \mathbb{Z}} h(L(\theta - m))$, which is periodic and localized in an interval of length $1/L$. We further normalize so that $\int_{-\infty}^{\infty} h(x)^2 dx = 1$, and hence $\int_0^1 h_L(\theta)^2 d\theta = 1/L$. Then set

$$(1) \qquad P(\theta) := \sum_{j=1}^{N} h_L(\theta - \theta_j)\langle \mathrm{Op}_N(f)\psi_j, \psi_j \rangle \ .$$
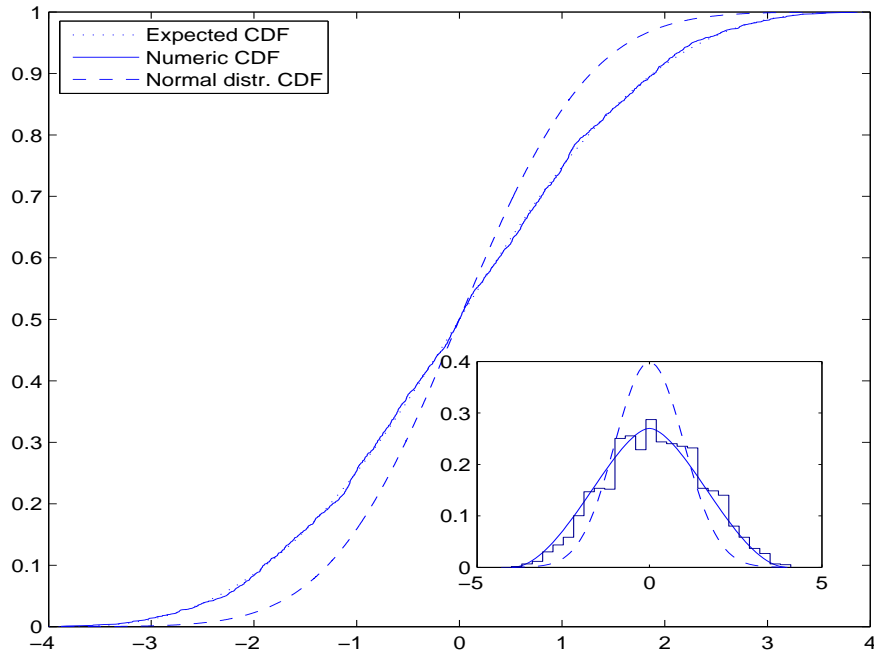
Figure 2. Cumulative distribution function of Hecke eigen-
basis compared to sum of two independent Sato-Tate's and
to a standard normal distribution

Note that $P(\theta)$ is independent of choice of basis, and in particular it is
real valued. An important case to consider is the case where $f(x)$ is a
trigonometrical function, we therefore denote for $n \in \mathbb{Z}^2$

$$P_n(\theta) := \sum_{j=1}^{N} h_L(\theta - \theta_j)\langle \mathrm{Op}_N(e(nx))\psi_j, \psi_j \rangle .$$

In [18] it was shown that if $\mathrm{ord}(A, N) \gg N^{1/2}$ then $\mathrm{Var}(\sqrt{L}P) \sim C(f) + o(1)$
where $C(f)$ is a constant depending on f and the matrix A. This variance
is the same variance as the limiting variance of the distribution of Hecke
matrix elements.

1.2. **Results.** In the following we present two results in the study of the
fluctuations of matrix elements. In section 7 we study the fluctuations in
short windows. In [18] we showed that unless n, m satisfy an arithmetic
condition, the corresponding fluctuation functions, $P_n(\theta), P_m(\theta)$, become
uncorrelated. In this paper we generalize this result for any choice of triple

$n_1, n_2, n_3 \in \mathbb{Z}^2$. That is we show that as $N \to \infty$ through primes

$$\int_0^1 P_{n_1}(\theta)P_{n_2}(\theta)P_{n_3}(\theta)d\theta = O(\frac{N}{L^{3/2}})$$

and in particular we prove the following theorem

Theorem 1. *Let $A \in SL_2(\mathbb{Z})$ be a hyperbolic matrix satisfying $A \equiv I$ (mod 2). Fix $f \in C^\infty(\mathbb{T}^2)$ of zero mean. Assume $L < 2\,\mathrm{ord}(A, N)$, then as $N \to \infty$ through split primes satisfying $\mathrm{ord}(A, N)/N^{2/3} \to \infty$, the third moment of $P(\theta)$ satisfies*

(2)
$$\int_0^1 \left(\sqrt{L}P(\theta)\right)^3 d\theta = o(1)$$

This results is consistent with a conjecture that $\sqrt{L}P$ has a Gaussian distribution (see section 9).

In section 8 we show agreement with the expected Sato-Tate limiting distribution and independent behaviour of the fluctuation of the martix elements for a fixed function. According to [22], the normalized matrix coefficient

$$\sqrt{N}\langle Op_N(f)\psi_j, \psi_j\rangle - \int_f$$

should be distributed like a weighted sum of traces of independent random matrices in $SU(2)$. In [22], the second and fourth moments are computed and shown to be consistent with this conjecture. We show that the fifth moment vanishes, in accordance to the conjecture:

Theorem 2. *Let $A \in SL_2(\mathbb{Z})$, $U_N(A)$ its quantization and $\{\psi_j\}_{j=1}^N$ a Hecke Basis. Fix $f \in C^\infty(\mathbb{T}^2)$. Then as $N \to \infty$ through primes,*

$$\frac{1}{N}\sum_{j=1}^N \left(\sqrt{N}(\langle Op_N(f))\psi_j, \psi_j\rangle - \int_{\mathbb{T}^2} f\right)^5 = O_f(\frac{1}{\sqrt{N}})$$

The results presented here are corollaries from bounds of mixed moments of a certain family of exponential sums. In sections 4,5,6 we introduce this family and study mixed moments of its distribution.

2. Background

2.1. Quantum mechanics on the torus.
We recall the basic facts of quantum mechanics on the torus which we need in the paper, see [28, 19] for further details. Planck's constant is restricted to be an inverse integer $1/N$, and the Hilbert space of states $\mathcal{H}_N$ is N-dimensional, which is identified with $L^2(\mathbb{Z}/N\mathbb{Z})$ with the inner product given by

$$\langle \phi, \psi \rangle = \frac{1}{N} \sum_{Q \bmod N} \phi(Q) \overline{\psi}(Q) .$$

Classical observables, that is real-valued functions $f \in C^\infty(\mathbb{T})$, give rise to quantum observables, that is self-adjoint operators $\mathrm{Op}_N(f)$ on $\mathcal{H}_N$. To define these, one starts with translation operators: For $n = (n_1, n_2) \in \mathbb{Z}^2$ let $T_N(n)$ be the unitary operator on $\mathcal{H}_N$ whose action on a wave-function $\psi \in \mathcal{H}_N$ is

$$T_N(n)\psi(Q) = e^{\frac{i\pi n_1 n_2}{N}} e(\frac{n_2 Q}{N})\psi(Q + n_1) .$$

For any smooth function $f \in C^\infty(\mathbb{T})$, define $\mathrm{Op}_N(f)$ by

$$\mathrm{Op}_N(f) = \sum_{n \in \mathbb{Z}^2} \widehat{f}(n) T_N(n)$$

where $\widehat{f}(n)$ are the Fourier coefficients of f. Below is a list of properties of $\mathrm{Op}_N(f)$:

(1) For $n = (n_1, n_2) \in \mathbb{Z}^2$, denote $\epsilon(n) = (-1)^{n_1 n_2}$, then

(3)
$$\mathrm{tr}(T_N(n)) = \begin{cases} \epsilon(n)N & n = 0 \pmod{N} \\ 0 & \text{otherwise} \end{cases}$$

and so for $f \in C^\infty(\mathbb{T}^2)$

$$\mathrm{tr}(\mathrm{Op}_N(f)) = N \int_{\mathbb{T}^2} f \, dx + O(\frac{1}{N^\infty})$$

(2) For $n, m \in \mathbb{Z}^2$

(4)
$$T_N(n)T_N(m) = e(\frac{\omega(m, n)}{2N})T_N(m + n)$$

2.2. Quantized cat map: Definition and results.
For $B \in SL_2(\mathbb{Z})$ the quantized cat $U_N(B)$ is a unitary operator on $\mathcal{H}_N$ satisfying "Exact Egorov" property

$$U_N(B)^* \mathrm{Op}_N(f) U_N(B) = \mathrm{Op}_N(f \circ B)$$

In [19] Kurlberg and Rudnick introduced a family of commuting operators $\mathcal{C}_A(N)$, called the Hecke group, which satisfy that after taking joint eigenfunctions of all elements in $\mathcal{C}_A(N)$, then all corresponding matrix elements satisfy

$$\left| \langle \mathrm{Op}_N(f)\psi, \psi \rangle - \int_{\mathbb{T}^2} f \right| \ll N^{-1/4-\epsilon}$$

and when N is restricted to primes, Gurevich and Hadani showed in [9] that the rate of convergence is in fact bounded by

$$|\langle \mathrm{Op}_N(f)\psi, \psi \rangle - \int_{\mathbb{T}^2} f| \le C(f) N^{-1/2}.$$

We restrict our discussion from now on to N prime. In this case, all but a finite subset of the primes A is diagonalizable over either $\mathbb{F}_N$ (the split case), or over $\mathbb{F}_{N^2}$ (the inert case). In the split case the group $\mathcal{C}_A(N)$ is isomorphic to $\mathbb{F}_N^*$, and in the inert case it is isomorphic to $\mathbb{F}_{N^2}^1$ the group of norm one elements in $\mathbb{F}_{N^2}$. In [22] Kurlberg and Rudnick exhibit some relations between Hecke matrix elements (matrix elements corresponding to Hecke eigenfunctions). For $A \in SL_2(\mathbb{Z})$ they introduced a quadratic form $Q(n)$ related to A

$$Q(n) = \omega(n, nA) = cn_1^2 + (d-a)n_1 n_2 - bn_2^2 \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

where $\omega(x, y) = x_1 y_2 - x_2 y_1$ is the standard symplectic form. As this quadratic form plays a crucial role in this paper, we list here some of its properties that were proven in [18], [22]:

(1) Since A is symplectic then A preserves $Q(n)$, that is for all $n \in \mathbb{Z}$ $Q(nA) = Q(n)$. Moreover, the Hecke group $\mathcal{C}_A(N)$ is isomorphic to $SO(Q, \mathbb{Z}/N\mathbb{Z})$. (§2.3 in [22])

(2) Let N be an odd prime, and let $A \in SL_2(\mathbb{Z})$ so that $(\mathrm{tr}\, A)^2 - 4 \ne 0 \mod N$. Then the space of binary quadratic forms preserved by A (which contains Q) is one dimensional. (Lemma 2.1 in [18])

(3) For $g \in \mathcal{C}_A(N)$, $g \ne 1$, denote by

$$q(x; g) := \omega(x(g-1)^{-1}, x(g-1)^{-1}g).$$

(Note that $q(\bullet, g) = 0$ if $g = -I \mod N$). Then if $g \ne \pm I \mod N$ then $q(\bullet; g)$ is a nonzero multiple of Q. Moreover,

(5) $$q(x; g) = \frac{Q(x)}{\lambda_A - \lambda_A^{-1}} \frac{1 + \lambda}{1 - \lambda}.$$

where λ_A is a generator of $\mathcal{C}_A(N)$ (§2.4 in [18])

Using this quadratic form, for a smooth function $f(x) \in C^\infty(\mathbb{T}^2)$ they introduce

$$f^\sharp(\nu) = \sum_{n: Q(n) = \nu} (-1)^{n_1 n_2} \hat{f}(n)$$

where $\hat{f}(n)$ are the Fourier coefficients of $f(x)$. They also conjecture the following

Conjecture 2.1. *Let $A \in SL_2(\mathbb{Z})$, $U_N(A)$ its quantization and $\{\psi_j\}_{j=1}^N$ a Hecke Basis. Let*

$$F_j^{(N)} = \sqrt{N} \left(\langle \mathrm{Op}_N(f)\psi_j, \psi_j \rangle - \int_{\mathbb{T}^2} f \right)$$

Then as $N \to \infty$ through primes, the limiting distribution of the normalized matrix elements $F_j^{(N)}$ is that of the random variable

$$X_f := \sum_{\nu \neq 0} f^\sharp(\nu) \operatorname{tr}(U_\nu)$$

where U_ν are independently chosen random matrices in $SU(2)$ endowed with Haar probability measure.

(See further discussion and properties in section 8.)

2.3. **Specific definition of $U_N(A)$.** In [15] Kelmer showed[1] that the following can be taken as a definition for $U_N(B)$

$$(6) \qquad U_N(B) = \frac{1}{N|\ker_N(B-I)|^{1/2}} \sum_{n \in (\mathbb{Z}/N\mathbb{Z})^2} e(\frac{\omega(n, nB)}{2N}) T_N(n(I-B))$$

where $\ker_N(B-I)$ denotes the kernel of the map $B-I$ on $\mathbb{Z}^2/N\mathbb{Z}^2$. We take this as the definition of $U_N(A)$ in this paper.

Proposition 2.2. Let $A \in SL_2(\mathbb{Z})$, $A \equiv I$ (mod 2). for any $B \in \mathcal{C}_A(N)$ let $U_N(B)$ be as in (6), then

(1) For any $I \neq B \in \mathcal{C}_A(N)$ $\operatorname{tr}(U_N(B)) = 1$
(2) Denote by

$$\varepsilon = \begin{cases} 1 & A \text{ inert} \\ -1 & A \text{ split} \end{cases}$$

and $\widetilde{U}_N(B) = \varepsilon U_N(B)$, then for $B_1, B_2 \in \mathcal{C}_A(N)$

$$(7) \qquad \widetilde{U}_N(B_1)\widetilde{U}_N(B_2) = \widetilde{U}_N(B_1 B_2)$$

Proof. (1) is an immediate result of the fact that

$$\operatorname{tr} T_N(n) = \begin{cases} N & n = 0 \pmod{N} \\ 0 & \text{otherwise} \end{cases}$$

and the fact that if $I \neq B \in \mathcal{C}_A(N)$ then $I - B$ is invertible. For (2) we recall that $U_N(A)$ is unique up to scalar multiplication, and there exist a choice $\widetilde{U}_N(A)$ that is multiplicative. In particular it was shown in [19],[21] that the eigenvalues of $\mathcal{C}_A(N)$ are characters of this group, and that they are all multiplicity free except the quadratic character, that in the inert case doesn't appear and in the split case appears with multiplicity 2. Therefore there exist a multiplicative choice of phase for which $\operatorname{tr}(\widetilde{U}_N(B)) = \varepsilon \chi_2(B)$. Since χ_2 is multiplicative we get that $\varepsilon \widetilde{U}_N(B) = U_N(B)$ is still multiplicative. $\square$

Remark 2.3. *From (7) we get that*

$$(8) \qquad U_N(A^t) = \varepsilon^{t-1} U_N(A)^t$$

---

[1]A similar formula for the p-adic metaplectique representation was shown already in [26]

2.4. **Fluctuations in short windows.** We recall in this section the basic setting from [18]. Denote by $h(t) = \mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]}(t)$ the characteristic function of the interval $[-\frac{1}{2}, \frac{1}{2}]$. Set

$$h_L(x) := \sum_{k \in \mathbb{Z}} h(L(x-k))$$

which is then a periodic function, localized on the scale of $1/L$, and $\int_0^1 h_L(\theta)^2 d\theta = 1/L$. The Fourier expansion of h_L is (in L^2 sense)

$$h_L(x) = \frac{1}{L} \sum_{t \in \mathbb{Z}} \widehat{h}\left(\frac{t}{L}\right) e(tx) .$$

where $\widehat{h}(y) = \int_{-\infty}^{\infty} h(x)e(-xy)\,dx$.

Let N be a prime which does not divide $\mathrm{disc}(Q) = (\mathrm{tr}\,A)^2 - 4$. Let

$$P(\theta) := \sum_j h_L(\theta - \theta_j)\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle$$

which is a sum of matrix elements on a window of size $1/L$ around θ. Then, in L^2 sense, and with $U = U_N(A)$, we have

$$(9) \qquad P(\theta) = \frac{1}{L} \sum_{t \in \mathbb{Z}} e(t\theta)\widehat{h}\left(\frac{t}{L}\right) \mathrm{tr}\{\mathrm{Op}_N(f)U^{-t}\} .$$

In [18] we proved the following results about $\mathrm{tr}\{\mathrm{Op}_N(f)U^{-t}\}$ (Lemma 2.3)

Lemma 2.4. *Let $A \in SL_2(\mathbb{Z})$ be hyperbolic, and assume that $A \equiv I \mod 2$. Then for any prime N not dividing $\mathrm{disc}(Q)$ and integer t such that $A^t \neq I \mod N$, we have*

$$(10) \qquad \mathrm{tr}\{T_N(k)U_N(A^t)\} = (-1)^{k_1 k_2} e\left(\frac{\overline{2}q(k; A^t)}{N}\right)$$

and in particular

$$\mathrm{tr}\{\mathrm{Op}_N(f)U_N(A^t)\} = \sum_k (-1)^{k_1 k_2} \widehat{f}(k)e\left(\frac{\overline{2}q(k; A^t)}{N}\right)$$

where $\overline{2}$ is the inverse of $2 \mod N$.

Following this lemma we denote by

$$(11) \qquad P_\nu(\theta) := \sum_j h_L(\theta - \theta_j)\langle T_N(n)\psi_j, \psi_j\rangle$$

where $\omega(n, nA) = \nu$ (this is well defined by lemma 2.4), and since $|\omega(n, nA)| \leq \|n\|_2^2 \|A\|_2^2$ we get the following decomposition of $P(\theta)$

Corollary 2.5. *Let* $f(x) = \sum_{\|n\| < R} \hat{f}(n) e(nx)$ *be a trigonometrical polynomial, then for* $N > \|A\|_2 R^2$

$$P(\theta) = \sum_{|\nu| < \|A\|_2 R^2} f^\sharp(\nu) P_\nu(\theta)$$

3. Background on exponential sums

We give some properties of exponential sums. For a given algebraic variety V over $k = \mathbb{F}_p$, and a given rational functions $f, g_1, \ldots, g_r$ on V defined over k, we denote $V(k)$ to be the k-rational points on V, and for multiplicative characters $\chi_1, \ldots, \chi_r$ and an additive character ψ, we define

$$(12) \qquad S = S(V, f) = \sum_{x \in V(k)} \psi(f(x)) \prod_{i=1}^{r} \chi(g_i(x))$$

and more generally, for any extension k_n of degree n of k, we define

$$(13) \qquad S_n = S(V \otimes_k k_n, f) = \sum_{x \in V(k_n)} \psi(\mathrm{tr}(f(x))) \prod_{i=1}^{r} \chi(N(g_i(x)))$$

where

$$
\begin{aligned}
N : k_n^* &\rightarrow k^* \\
x &\mapsto x \cdot x^p \ldots x^{p^{n-1}} \\
\mathrm{tr} : k_n &\rightarrow k \\
x &\mapsto x + x^p + \cdots + x^{p^{n-1}}
\end{aligned}
$$

are the norm and trace maps respectively. All these sums are packaged in the corresponding L - function:

$$(14) \qquad L(S, T) = \exp\left(\sum_{n=1}^{\infty} \frac{S_n}{n} T^n\right)$$

3.1. **Deligne's results.** The following was proven by Deligne in [5]:
   (1) $L(S, T)$ is a rational function.
   (2) The exponential sums S_n satisfy:

$$(15) \qquad S_n = \sum_i \alpha_i^n - \sum_i \beta_i^n$$

   where α_i are the inverse of the zeros of $L(S, T)$, and β_i are inverse of its poles, and both are called the roots of the exponential sum.
   (3) The roots are algebraic integers.
   (4) All conjugates of a root have the same absolute value which is a positive integer power of $\sqrt{p}$.

It was proved by Katz in [13], that there exists a constant C, independent of p, such that for a given exponential sum of type (12) there are at most C roots.

3.2. **Weil's results.** For a 1 dimensional exponential sum, there is no need for the full power of Deligne's work, but rather the proof of Weil for RH over finite fields. We state below the main results concerning this paper.

(1) Let $\mathbb{F}$ be finite field of q elements, and $\mathbb{F}[x]$ the ring of polynomials over $\mathbb{F}$. For a polynomial $Q(x) \in \mathbb{F}[x]$, and a multiplicative character modulo Q, we define the corresponding L - function

$$L(u, \chi) = \prod_{P|Q} \left(1 - \chi(P)u^{\deg P}\right)^{-1}$$

where the product is over all irreducible monic polynomials in $\mathbb{F}[x]$. By unique factorization in $\mathbb{F}[x]$, we have that

$$L(u, \chi) = \sum_{f \neq 0} \chi(f)u^{\deg f} = \sum_{n=1}^{\infty} a_n(\chi)u^n$$

where the sum is over all monic polynomials in $\mathbb{F}[x]$, and $a_n(\chi) = \sum_{\deg f = n} \chi(f)$.

(2) For a nontrivial character modulo Q, $L(u, \chi)$ is in fact a polynomial in u, of degree at most $\deg Q - 1$. We may factor it as follows

$$L(u, \chi) = \prod_{j=1}^{\deg Q - 1} (1 - \alpha_j(\chi)u)$$

and it was shown by Weil [31], that for all $j = 1, \ldots, \deg L(u, \chi)$, $|\alpha_j(\chi)| \leq \sqrt{q}$. Note also that $a_1(\chi) = -\sum_{j=1}^{\deg Q - 1} \alpha_j(\chi)$.

3.3. **Bound for double exponential sums.** In this part we prove a lemma that gives a sufficient condition for an exponential sum to have square root cancelation.

We first prove the following proposition that appears previously in [1, 12].

Proposition 3.1. *Let $\xi_1, \ldots, \xi_n \in \mathbb{C}$ be distinct complex numbers of absolute value one, and $b_1, \ldots, b_n \in \mathbb{C}$ complex numbers. Then*

$$\limsup_{\nu \to \infty} |\sum_{i=1}^{n} b_1 \xi_i^\nu| \geq \left(\sum_{i=1}^{n} |b_i|^2\right)^{1/2}$$

Proof. For $N \in \mathbb{N}$ compute the average over $\nu = 1, \ldots, N$
(16)
$$\frac{1}{N}\sum_{\nu=0}^{N-1} |\sum_{i=1}^{n} b_i \xi_i|^2 = \sum_{i=1}^{n} |b_i|^2 + \frac{1}{N} \sum_{i \neq j} b_i \overline{b_j} \frac{1 - (\xi_i \xi_j^{-1})^N}{1 - (\xi_i \xi_j^{-1})} = \sum_{i=1}^{n} |b_i|^2 + O(\frac{1}{N})$$

Now assume that

$$\limsup_{\nu \to \infty} |\sum_{i=1}^{n} b_i \xi_i^\nu|^2 < \sum_{i=1}^{n} |b_i|^2 - \delta$$

for some $\delta > 0$. Then in particular the bound is true for ν large enough, which contradicts (16). $\square$

The next lemma shows that for general exponential sums, given a bound on the sum of squares can can lead to a bound on individuals.

Lemma 3.2. *Let k be a finite field of characteristic $char k = p$, and V be an algebraic variety over $\overline{k}$ of dimension N and degree d. Let $\underline{\chi} = \{\chi_1, \ldots, \chi_l\}$ be muultiplicative characters of k^*, and ψ an and additive character of k, $g_1(x), \ldots, g_l(x), f(x)$ rational functions over V. Denote*

$$S(\chi, \psi; \underline{g}, f) := \sum_{x \in V(k)} \prod_{i=1}^{l} \chi_i(g_i(x))\psi(f(x))$$

Assume that there exists $b \in \mathbb{N}$ and $M \in \mathbb{R}$ such that for all $\nu \in \mathbb{N}$

$$(17) \qquad \frac{1}{|k|^{\nu}} \sum_{0 \neq a \in k_{\nu}} \left| S_{\nu}(\underline{\chi}, \psi_a; \underline{g}, f) \right|^2 \leq M|k|^{\nu b}$$

where $\psi_a(x) = \psi(\mathrm{tr}(ax))$, then there exists a constant $B = B(N, d, f, g)$ such that $|S(\underline{\chi}, \psi; \underline{g}, f)| \leq B|k|^{b/2}$

Proof. By Deligne's result, it suffice to show that all the roots ω_i of (the L-function of) the exponential sum $S(\chi, \psi_a)$ are of absolute value $|\omega_i| \leq |k|^{b/2}$. Denote $r_{max} = \max_{w_i \text{ roots of } S}\{r : |w_i| = |k|^{r/2}\}$, and assume $r_{max} > b$ so there exist $\omega_1, \ldots, \omega_n$ roots of $S(\underline{\chi}, \psi; \underline{g}, f)$, of absolute value $|k|^{r_{max}/2}$, and multiplicities $\lambda_1, \ldots, \lambda_n$, then by proposition 3.1 with $\xi_i = w_i/|k|^{r_{max}/2}, b_i = \lambda_i$ we get

$$\limsup_{\nu \to \infty} \left| \frac{\sum_{i=1}^{n} \lambda_i \omega_i}{|k|^{r_{max}\nu/2}} \right| > 0$$

and therefore there exist infinitely many ν_j such that $|S_{\nu_j}(\underline{\chi}, \psi; \underline{g}, f)| \gtrsim |k|^{r_{max}\nu_j/2}$. for $0 \neq c \in \{1, \ldots, p-1\}$, let $\sigma_c \in Gal(\mathbb{Q})$ that sends $e(1/p) \mapsto e(c/p)$. Since the fields $\mathbb{Q}(e(1/p)), \mathbb{Q}(e(1/(p-1)))$ are linearly disjoint we get that $S(\underline{\chi}, \psi; \underline{g}, f)^{\sigma} = S(\underline{\chi}, \psi_c; \underline{g}, f)$, and therefore

$$\sum_{0 \neq a \in k_{\nu_j}} |S_{\nu}(\underline{\chi}, \psi_a; \underline{g}, f)|^2 \geq \sum_{0 \neq a \in k} \left| S_{\nu}(\underline{\chi}, \psi_a; \underline{g}, f) \right|^2 \gtrsim |k|^{r_{max}\nu_j + 1}$$

which contradicts (17) $\square$

For exponential sums over a 2 dimensional variety the following theorem gives a sufficient condition for square root cancelation.

Theorem 3.3. *Let V be an irreducible algebraic variety over a finite field k of dimension 2, and degree δ. Let f be rational function on V. Suppose that there exists R such that*

$$\#\{C \in \overline{k} : \text{the fiber f = C is geometrically reducible}\} < R$$

and that the degree of all irreducible fibers is at most d. Let ψ be an additive character of k. Then there exists B such that

$$(18) \qquad |\sum_{x \in V(k)} \psi(f(x))| \leq B|k|$$

Proof. By lemma 3.2 it suffices to show that there exists M such that

$$\frac{1}{|k|^\nu} \sum_{0 \neq a \in k_\nu} |S(\psi_a)|^2 \leq M|k|^{2\nu}$$

which is equivalent to show that

$$(19) \qquad \sum_{a \in k_\nu} |S(\psi_a)|^2 = N_\nu^2 + O(|k|^{3\nu})$$

where $N_\nu = \#\{\vec{x} \in V(k_\nu)\} = |k|^{2\nu} + O(|k|^{3\nu/2})$ by irreducibility of V and Lang-Weil theorem [23]. Writing the sum in (19) explicitly we get

$$\sum_{a \in k_\nu} \sum_{x \in V(k_\nu)} \sum_{x' \in V(k_\nu)} \psi(a(f(x) - f(x'))) =$$

$$|k|^\nu \sum_{x \in V(k_\nu)} \sum_{x' \in V(k_\nu)} \sum_{f(x)=f(x')} 1 = |k|^\nu \sum_{C \in k_\nu} |f_\nu^{-1}(C)|^2$$

Where $f_\nu^{-1}(C) = \{x \in V(k_\nu) : f(x) = C\}$. The number of points on $f_\nu^{-1}(C)$ is given by

$$(20) \qquad |f_\nu^{-1}(C)| = \frac{1}{|k|^\nu} \sum_{a \in k_\nu} \sum_{\vec{x} \in V(k_\nu)} \psi(a(f(\vec{x}) - C)) = \frac{N_\nu}{|k|^\nu} + E_\nu(C)$$

where

$$(21) \qquad E_\nu(C) = \frac{1}{|k|^\nu} \sum_{0 \neq a \in k_\nu} \sum_{\vec{x} \in V(k_\nu)} \psi(a(f(\vec{x}) - C))$$

and therefore

$$|k|^\nu \sum_{C \in k_\nu} |f_\nu^{-1}(C)|^2 = |k|^\nu \sum_{C \in k_\nu} \left(\frac{N_\nu}{|k|^\nu} + E_\nu(C)\right)^2 =$$

$$N_\nu^2 + \sum_{C \in k_\nu} \left(2N_\nu E_\nu(C) + |k|^\nu E_\nu(C)^2\right)$$

By the assumption on the fibers and the Riemann hypothesis for curves we get that $E_\nu(C) = O(|k|^{\nu/2})$ for all C except at most R. From (21) we have that $\sum_{C \in k_\nu} \tilde{E}_\nu(C) = 0$ and therefore we get that

$$\sum_{a \in k_\nu} |S(\psi)|^2 = N_\nu^2 + O(|k|^{3\nu})$$

which concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Remark 3.4. *As was seen throughout the proof the irreducibility assumptions can be replaced by cardinality assumptions on V and the fibers, that is if $\#V = |k|^2 + O(|k|^{3/2})$ and the fibers satisfy $|f_\nu^{-1}(C) - |k|^\nu| \leq B\sqrt{|k|^\nu}$ for an absolute constant B then the theorem holds as well.*

3.4. **Bounds for character sums over $\mathbb{F}_q$.** We prove here a condition for a square root cancelation for one dimensional sums involving many multiplicative characters. We prove the following

Theorem 3.5. *Let $k = \mathbb{F}_q$ be the field with $q = p^n$ elements ($char(k) = p$), and let $\chi_1, \ldots, \chi_m$ be nontrivial multiplicative characters of k. Let $P_1(x), \ldots, P_m(x) \in k[x]$ be monic irreducible polynomials of degrees $d_1, \ldots, d_m$ respectively. Then*

$$(22) \qquad \sum_{t \in \mathbb{F}_q} \prod_{i=1}^m \chi_i(P_i(t)) \leq \left(\sum_{i=1}^m d_i - 1 \right) \sqrt{q}$$

To prove this bound we construct a polynomial $Q_\chi(x) \in k[x]$ of degree less then $\sum d_i$, and a nontrivial character $\nu_\chi : (k[x]/Q_\chi(x))^x \to \mathbb{C}$, such that $\nu_\chi(x - t) = \prod_{i=1}^m \chi_i(P(t))$.

Proposition 3.6. *Let k be a field, and let $\{x_1, \ldots, x_l\} \subset \bar{k}$ be finite set invariant under Galois action. Then for any set $y = \{y_1, \ldots, y_l\} \subset \bar{k}$ invariant under Galois action, there exists a unique monic polynomial $P_y(x) \in k[x]$ of degree l, such that $P_y(x_i) = a_i, \forall \sigma \in Gal(k) \; P_y(\sigma(x_i)) = \sigma(a_i)$.*

Proof. The existence and uniqueness of $P(x) \in \bar{k}[x]$ is a standard linear algebra argument. To show that $P(x) \in k[x]$ we notice that for any $\sigma \in Gal(k)$

$$\sigma(P)(x_i) = \sigma(P(\sigma^{-1}(x_i))) = \sigma(\sigma^{-1}(P(x_i))) = P(x_i)$$

and by uniqueness of $P(x)$ we get that $\sigma(P)(x) = P(x)$ and hence $P(x) \in k[x]$. $\qquad\square$

Proposition 3.6 will give us a way to construct the required ν_χ. We do it using the resultant of two polynomials

Definition 3.7. *Let $P, Q \in k[x]$ be two monic polynomials. Define*

$$res(P, Q) := \prod_{\substack{x_i \text{roots of } Q \\ y_j \text{roots of } P}} (x_i - y_j) = \prod_{y_j \text{roots of P}} Q(y_j) = (-1)^{deg(P)} \prod_{x_i \text{roots of Q}} Q(x_i)$$

Corollary 3.8. *Let $P_1, \ldots, P_m \in k[x]$ be distinct monic irreducible polynomials of degrees $d_1, \ldots, d_m$ respectively. Then for any $a = (a_1, \ldots, a_m) \in k^m$ there exists a polynomial $Q_a(x) \in k[x]$ such that $res(P_i, Q_a) = a_i$*

Proof. For any P_i let $Y_i = \{y_{ij}\}_{j=1}^{d_i} \subset \bar{k}$ be the set of its roots in the algebraic closure. Then $Y = \cup_i Y_i$ is an invariant set under the Galois group action. For a_i let $Z_i = \{z_{i1}, \ldots, z_{id_i}\}$ be a set of Galois conjugates elements in $k_i := k(y_{i1}, \ldots, y_{id_i})$ such that $N_{k_i/k}(z_i) = z_1 \cdots z_{d_i} = a_i$, and

let $Z = \cup_i Z_i$. By proposition 3.6 there exists a polynomial $Q_a(x) \in k[x]$ such that $Q_a(y_{ij}) = z_{ij}$. Then

$$res(P_i, Q_a) = \prod_{y_{ij} \text{roots of } P} Q_a(y_{ij}) = z_{i1} \cdots z_{id_i} = a_i$$

$\square$

We can now conclude the proof of theorem 3.5. Denote by $Q(x) = lcm(P_1, \ldots, P_m) \in k[x]$, and by ν_χ the character of $(k[x]/Q(x))^x$ defined by

$$\nu_\chi(F) = \prod_{i=1}^{m} \chi_i(res(P_i, F))$$

By previous corollary ν_χ is nontrivial, and by definition of the resultant it is well defined modulo Q. Thus by Weil's result the theorem is proved.

4. A FAMILY OF EXPONENTIAL SUMS

Let k be a finite field of $q = p^n$ elements. For ψ be an additive character of k and χ a multiplicative character of k^* We define the following exponential sum

(23) $$F(\chi; \psi) = \sum_{0,1 \neq x} \chi(x) \psi(\frac{1+x}{1-x})$$

We consider the family $\{F(\chi; \psi)\}_\chi$ where χ runs through all characters of k^*. It was shown in [18] that

(24) $$|F(\chi; \psi)| \leq \sqrt{q}$$

In light of this result we normalize the sum and define

$$\tilde{F}(\chi; \psi) = \frac{F(\chi; \psi)}{\sqrt{q}}$$

and consider the family $\{\tilde{F}(\chi; \psi)\}_{\chi, \psi}$. The following proposition give some basic properties of this family

Proposition 4.1. *Let $F(\chi; \psi)$ be as above.*

(1) *For any pair χ, ψ as above, the sum $F(\chi; \psi)$ is r a real number.*

(2) $$\frac{1}{q-1} \sum_\chi \tilde{F}(\chi; \psi) = 0$$

(3) $$\frac{1}{q-1} \sum_\chi |\tilde{F}(\chi; \psi)|^2 = 1 - \frac{2}{q}$$

(4) *For any $\chi_1 \neq \chi_2$, and ψ_1, ψ_2*

(25) $$\frac{1}{q-1} \sum_\chi \tilde{F}(\chi_1 \chi; \psi_1) \overline{\tilde{F}(\chi_2 \chi; \psi_2)} \ll \frac{1}{\sqrt{q}}$$

Proof. The first part of the proposition follows from the simple observation that

$$\overline{F(\chi;\psi)} = \sum_{0,1\neq x} \chi(x^{-1})\psi(-\frac{1+x}{1-x}) = \sum_{0,1\neq x} \chi(x^{-1})\psi(\frac{1+x^{-1}}{1-x^{-1}}) = F(\chi;\psi)$$

Parts 2,3 of the proposition are immediate consequence of the orthogonality relations of characters. For (25) we have

(26)
$$\frac{1}{q-1}\sum_{\chi} \frac{F(\chi_1\chi;\psi_1)}{\sqrt{q}} \frac{\overline{F(\chi_2\chi;\psi_2)}}{\sqrt{q}} = \frac{1}{q}\sum_{0,1\neq x} \chi_1\overline{\chi_2}(x)\psi_1\overline{\psi_2}(\frac{1+x}{1-x}) \ll \frac{1}{\sqrt{q}}$$

where

$$\chi_1\overline{\chi}_2(x) = \chi_1(x)\chi_2(x^{-1})$$
$$\psi_1\overline{\psi}_2(y) = \psi_1(y)\psi_2(-y)$$

ans the last inequality is due to (24). $\square$

The last proposition can be considered as computation of mean and variance for fixed ψ and running over χ, and the third result as covariance for two random variables $\tilde{F}(\chi_1\chi;\psi), \tilde{F}(\chi_2\chi;\psi)$ running over χ, and in fact proves that for any two additive characters ψ_1,ψ_2 the random variables $\tilde{F}(\chi;\psi_1), \tilde{F}(\chi;\psi_2)$ become uncorrelated. The following conjecture suggests even a stronger behaviour.

Conjecture 4.2. *Let $\chi,\psi,F(\chi;\psi)$ be as defined above, then as $q \to \infty$ through primes, we have the following:*

(1) *The sets $\{\frac{F(\chi;\psi)}{\sqrt{N}}\}_\chi$ become equidistributed with respect to the Sato-Tate distribution μ_{ST}, that is the distribution of $\mathrm{tr}(U)$ where $U \in SU(2)$ is random matrix with respect to Haar measure.*

(2) *For any finite field k, let $(\chi_i,\psi_i)_{i=1}^m$ be a set of m distinct pairs of multiplicative and additive characters of k. Then the sets*

$$\{(\frac{\tilde{F}(\chi_1\chi;\psi_1)}{\sqrt{q}}, \frac{\tilde{F}(\chi_m,\psi_m)}{\sqrt{q}})\}_\chi$$

become equidistributed with respect to the product of m Sato-Tate measures, that is they become independent.

(3) *In particular, the mixed moments of m distinct pairs*

$$(\tilde{F}(\chi_i;\psi_i),\ldots,\tilde{F}(\chi_m;\psi_m))$$

satisfy

(27)
$$\frac{1}{|q-1|}\sum_{\chi}\prod_{i=1}^m \left(\tilde{F}(\psi_i;\chi\chi_i)\right)^{e_i} = \mathbb{E}(\prod_{i=1}^m X_{\psi_i,\chi_i}^{e_i}) + O(\frac{1}{\sqrt{q}})$$

where X_{ψ_i,χ_i} $i = 1,\ldots,m$ are IID random variables with Sato-Tate distribution

In the following sections we give some agreement with this conjecture by proving the following theorems

Theorem 4.3. *Let k be a finite field with $q = p^n$ elements, and let χ_1, χ_2, χ_3 be any 3 multiplicative characters of k^*, and ψ_1, ψ_2, ψ_3 be any nontrivial additive characters of k, then*

$$(28) \qquad \frac{1}{q-1} \sum_\chi \tilde{F}(\chi_1\chi; \psi_1) \tilde{F}(\chi_2\chi; \psi_2) \tilde{F}(\chi_3\chi; \psi_3) \ll \frac{1}{\sqrt{q}}$$

Theorem 4.4. *Let k be a finite field with $q = p^n$ elements, let χ_1 be any multiplicative character of k^*, and $\psi_1, \ldots, \psi_5$ be any nontrivial additive characters of k, then*

$$(29) \qquad \frac{1}{q-1} \sum_\chi \prod_{i=1}^5 \tilde{F}(\chi_1\chi; \psi_i) \ll \frac{1}{\sqrt{q}}$$

Remark 4.5. *In section 9 we show numerical evidence for conjecture 4.2.*

5. Proof of theorem 4.3

We start by making a change of variables in the sum over χ by letting $\chi \mapsto \overline{\chi}_3\chi$, and we therefore may assume that χ_3 is trivial. Moreover since ψ_1, ψ_2, ψ_3 are nontrivial, there exists $0 \neq A_0, A, B \in k$ such that $\psi_1(x) = \psi(A_0 x), \psi_2(x) = \psi(Ax), \psi_3(x) = \psi(Bx)$ where ψ is a generator of the group of additive characters. We next sum over χ to get

$$(30) \quad S(\chi_1, \chi_2; \psi) := \sum_{x,y} \chi_1(x)\chi_2(y)\psi\left(A_0\frac{1+x}{1-x} + A\frac{1+y}{1-y} + B\frac{xy+1}{xy-1}\right)$$

(without loss of generality we assume $\psi = \psi_1$ and therefore $A_0 = 1$). Under the change of variables $(x, y) \mapsto (\frac{1-x}{1+x}, \frac{1-y}{1+y})$, the sum changes to
(31)

$$S(\chi_1, \chi_2; \psi) = \sum_{x,y,xy\neq 0,-1} \chi_1\left(\frac{1-x}{1+x}\right)\chi_2\left(\frac{1-y}{1+y}\right)\psi\left(x + Ay - B\frac{xy+1}{x+y}\right)$$

Proposition 5.1. *Let k be a finite field. For $0 \neq A, B \in k$ let $f(x, y) = x + Ay - B\frac{xy+1}{x+y}$. Then for all $C \in \overline{k}$ satisfying $C^2 \neq (A - B - 1)^2 - 4B$ the fiber $f(x, y) = C$ is absolutely irreducible*

Proof. For $C \in \overline{k}$ consider the equation

$$f(x, y) = x + Ay - B\frac{xy+1}{x+y} = C$$

multiplying it by $x + y$ this turns out to be

$$x^2 + (A - B + 1)xy + y^2 - Cx - Cy - B = 0$$

which is a quadratic curve. For a quadratic curve $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c$ it is known that it is absolutely irreducible over k if the determinant

$$\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{vmatrix} \neq 0$$

In our case it is $\frac{B}{4}\left(C^2 - (A^2 + B^2 - 2A - 2B - 2AB + 1)\right)$, since $B \neq 0$ we get that for at most 2 values of C the curve is reducible. Furthermore, since it is an irreducible quadratic curve, its intersection with the curve $x + y = 0$ has at most 2 affine points points and therefore multiplying by that factor was valid. $\qquad\square$

Corollary 5.2. *Let χ_1 be any multiplicative character of k. Then there exists $0 < M \in \mathbb{R}$ such that*

$$(32) \qquad \frac{1}{q-1}\sum_\chi \tilde{F}(\chi_1\chi;\psi_1)\tilde{F}(\chi_1\chi;\psi_2)\tilde{F}(\chi_1\chi;\psi_3) \leq M\frac{1}{\sqrt{q}}$$

Proof. This is an immediate corollary of theorem 3.3 $\qquad\square$

For the cases were not all characters are equal we use the following proposition that observes some geometric properties of the fibers.

Proposition 5.3. *Let k, $f(x,y)$ be as above, and $C \in \overline{k}$ satisfying $C^2 \neq (A - B - 1)^2 - 4B, 0$. Denote $f_C = \{(x,y) \in k^2 : f(x,y) = C\}$. Then f_C satisfies*

(1) *For any $a \in k$ the intersection of the curve $x - a$ or $y - a$ with f_C has at most 2 points.*

(2) *For any $a \in k$, the intersection of the curve $\frac{x-1}{x+1} = a\frac{y-1}{y+1}$ with f_C has at most 4 points.*

(3) *For any $a \in k$, the intersection of the curve $\frac{x-1}{x+1} = a\frac{y+1}{y-1}$ with f_C has at most 4 points.*

(4) *The intersection of the curve $(x^2 - 1)(y^2 - 1) = 0$ with f_C has at most 8 points.*

(5) *If f_C is not empty, neither $\frac{x+1}{x-1}$ nor $\frac{y+1}{y-1}$ are contained in any multiplicative coset of the subgroup of squares of k^* along f_C.*

Proof. By assumption on C, and by proposition 5.1 we have that f_C is an irreducible quadratic curve. As such, if G is any other curve (not necessarily irreducible) not containing f_C, the number of points $\sharp G \cap f_C$ is bounded by the product $deg(G) \cdot deg(f_C) = 2deg(G)$. Therefore, since $x - a, y - a, (1 - a)(xy - 1) - (1 + a)(x - y), (1 - a)(xy + 1) + (1 + a)(x + y), (x^2 - 1)(y^2 - 1)$ are all coprime to $f(x,y) - C$ we get properties 1, 2, 3, 4. For property 5 we give a parametrization of f_C. Choose a point $(x_1, y_1) \in f_C$ such that $x_1 \neq 0, \pm1$ (such a choice is possible since f_C is not empty and by property

1). Then the following is a parametrization of f_C:

$$(33) \qquad x(t) = \frac{x_1(At^2 - Ct - B)}{At^2 - ((A - B + 1)x_1 + 2Ay_1)t + C(x_1 + y_1) + B}$$

$$(34) \quad y(t) = \frac{(C - (A - B + 1)x_1 - Ay_1)t^2 + t(2B + Cx_1) - By_1}{At^2 - t((A - B + 1)x_1 - 2Ay_1) + C(x_1 + y_1) + B}$$

This gives similar expressions for $\frac{x-1}{x+1}, \frac{y-1}{y+1}$. Denote by $q_1(t), q_2(t)$ the numerator and denominator for $\frac{x-1}{x+1}$ and $q_3(t), q_4(t)$ the numerator and denominator for $\frac{y-1}{y+1}$. Direct computation gives that

$$disc(q_1(t)) = disc(q_3(t)) = (C - A + B - 1)^2 x_1^2$$
$$disc(q_2(t)) = disc(q_4(t)) = (C + A - B + 1)^2 x_1^2$$

Therefore the only possibility for all to be squares is if $A - B + 1 = 0, C = 0$, by choosing $C \neq 0$ we get property 5. $\qquad \square$

Corollary 5.4. *Let k be a finite field with $q = p^n$ elements, and $f(x,y)$ as above. Let χ_1, χ_2 be multiplicative characters of k not both trivial. For $\nu \in \mathbb{N}$ be the extension of k of degree ν, Then for all $C \in k_\nu$ satisfying $C^2 \neq (A - B - 1)^2 - 4B, 0$ we have*

$$(35) \qquad | \sum_{f(x,y)=C} \chi_1(\frac{1-x}{1+x})\chi_2(\frac{1-y}{1+y})| \leq 7\sqrt{q}$$

Proof. By proposition 5.1 the sets $f_C = \{(x,y) \in k^2 : f(x,y) = C\}$ are irreducible quadratic curves, and hence as in proposition 5.3 we can parameterize the curve and the sum becomes

$$(36) \quad \sum_{t\in k}\chi_1(q_1(t))\chi_2(q_2(t)) = \sum_{t\in k}\chi_1(p_1(t))\overline{\chi}_1(p_2(t))\chi_2(p_3(t))\overline{\chi}_2(p_4(t))$$

where

$$q_1(t) = \frac{p_1(t)}{p_2(t)}, q_2(t) = \frac{p_3(t)}{p_4(t)}, i = 1, 2$$

This can be written in the form

$$\varepsilon \sum_{t\in k}\prod_{i=1}^{m} \widetilde{\chi}_i(\tilde{p}_i(t))$$

where $\tilde{p}_i(t)$ area different monic irreducible polynomials, $|\varepsilon| = 1$, by decomposing $p_i(t), i = 1, \ldots, 4$ into irreducible parts, joining equal parts together, and powers are absorbed into the characters. In order to apply theorem 3.5, we must check that there exists a nontrivial pair $\widetilde{\chi}_i, \tilde{p}_i(t)$, that is, there exists $1 \leq i \leq m$, such that $\widetilde{\chi}_i$ is not the trivial character, and $\tilde{p}_i(t) \not\equiv 1$. However $\widetilde{\chi}_i, i = 1, \ldots, 4$ are the trivial characters, and $\tilde{p}_i(t), i = 1, \ldots, 4$ are constant if complete cancelation occurs already in

$$\chi_1(p_1(t))\overline{\chi}_1(p_2(t))\chi_2(p_3(t))\overline{\chi}_2(p_4(t))$$

that is one of the following holds

(1) $p_1(t)/p_2(t) = Const., p_3(t)/p_4(t) = Const$
(2) $p_1(t)/p_3(t) = Const., p_2(t)/p_4(t) = Const$, and $\chi_1 = \overline{\chi}_2$
(3) $p_1(t)/p_4(t) = Const., p_2(t)/p_3(t) = Const$, and $\chi_1 = \chi_2$
(4) $p_1 p_2/(p_3 p_4) = Const.$, and $\chi_1^2 = \chi_2^2$ are trivial.
(5) $p_1, p_2, p_3, p_4 = \square$, and $\chi_1^2 = \chi_2^2$ are trivial.

and each of these conditions corresponds to a geometric restriction that was proved impossible in proposition 5.3. We can therefore imply theorem 3.5. Since $\deg(p_i(t)) \leq 2$, then the sum of their degrees is at most 8, hence the sum of the degrees of $\widetilde{p}_i(t)$ is at most 8 also, and therefore we get the required result.   □

We can now conclude the proof of theorem 4.3. Let χ_1, χ_2 be multiplicative characters of k not both trivial, we want to show that there exists $M > 0$ such that the following bound holds

(37) $$|S(\chi_1, \chi_2; \psi)| \leq M|k|$$

where $S(\chi_1, \chi_2; \psi)$ was defined in (30).

Proposition 5.5. *Let $S(\chi_1, \chi_2; \psi)$ be as above. Then for all $\nu \in \mathbb{N}$*

$$\frac{1}{|k_\nu|} \sum_{a \in k_\nu} |S(\chi_1 \circ N, \chi_2 \circ N; (\psi \circ tr)_a)|^2 \leq 52|k|^{2\nu}$$

Proof. To show the bound we compute the sum over a to get

$$\frac{1}{|k_\nu|} \sum_{a \in k_\nu} |S(\chi_1 \circ N, \chi_2 \circ N; (\psi \circ \mathrm{tr})_a)|^2 =$$

$$\sum_{C \in k_\nu} | \sum_{x,y \in k_\nu : f(x,y)=C} \chi_1 \circ N(\frac{1-x}{1+x})\chi_2 \circ N(\frac{1-y}{1+y})|^2$$

and by corollary 5.4 the inner sum is bounded by $49|k_\nu|$ for all but at most 3 values of C, hence the bound is proved.   □

Corollary 5.6. *Let k be a finite field with $char(k) \neq 2$. Then there exists $M > 0$ such that for any two multiplicative characters, and ψ additive character of k the sum $S(\chi_1, \chi_2; \psi)$ satisfies*

$$|S(\chi_1, \chi_2; \psi)| \leq M|k|$$

Proof. This is an immediate corollary of proposition 5.5 and lemma 3.2   □

6. PROOF OF THEOREM 4.4

For the proof of theorem 4.4 we use an averaging technique that will allow us to distinguish symmetries of the sum.

Definition 6.1. *Let k be a finite field, χ_1, χ_2 be multiplicative characters of k, ψ an additive character and $a \in k^*$. Denote*

$$F(\chi_1, \chi_2; \psi, a) = \sum_{x \in k} \chi_1(x+a)\overline{\chi}_2(x-a)\psi(a^{-1}x)$$

$$G(\chi_1, \chi_2; \psi) = \frac{1}{|k^*|} \sum_{a \in k^*} F(\chi_1, \chi_2; \psi, a)$$

Proposition 6.2. *Let k be a finite field, χ_1, χ_2, ψ, a as above. Then*

   (1) $F(\chi_1, \chi_2; \psi, a) = \chi_1\overline{\chi}_2(a)F(\chi_1, \chi_2; \psi, 1)$

   (2)

(38)
$$G(\chi_1, \chi_2; \psi) = \begin{cases} F(\chi_1; \psi) & \chi_1 = \chi_2 \\ 0 & \chi_1 \neq \chi_2 \end{cases}$$

   (3) *Let θ be any multiplicative character of k, and $\psi_1, \ldots, \psi_5$ any non-trivial additive characters of k, then*

(39)
$$\frac{1}{|k^*|} \sum_{\chi} \prod_{i=1}^{5} F(\theta\chi; \psi_i) = \frac{1}{|k^*|^6} \sum_{\chi_1,\ldots,\chi_5 \in \widehat{k^*}} \prod_{i=1}^{5} G(\chi_i, \chi_{i+1}; \psi_i)$$

where inside the product we consider 5+1 as 1.

Proof. Part 1 is immediate under the change of variables $x \mapsto ax$ which is invertible since $a \in k^*$. Part 2 and the equality in (39) are then immediate from part 1 and the orthogonality relations. $\qquad\square$

Lemma 6.3. *Let θ be a multiplicative character of k ($char(k) \neq 2$), ψ a nontrivial additive character of k. for $A_1, \ldots, A_5 \in k^*$ let $\psi_i(x) = \psi(A_i x)$ be nontrivial additive characters of k. Denote by*

$$\mathbf{V}(\underline{A}) = \left\{ 0 \neq a_1, a_2, a_3, a_4 : \begin{smallmatrix} A_1 a_1 + A_2 a_2 + A_3 a_3 + A_4 a_4 + A_5 = 0 \\ a_1^{-1} + a_2^{-1} + a_3^{-1} + a_4^{-1} + 1 = 0 \end{smallmatrix} \right\}$$

and

$$\tilde{h}_A(a_2, a_3, a_4) = \sum_{2 \leq i < j \leq 4} \left(\frac{A_i a_i}{a_j} - \frac{A_j a_j}{a_i} \right) + \sum_{i=1}^{4} A_i a_i - A_5 \sum_{i=2}^{4} a_i^{-1}$$

then

(40)
$$\frac{1}{|k^*|} \sum_{\chi} \prod_{i=1}^{5} \tilde{F}(\theta\chi; \psi_i) =$$

$$\frac{1}{|k|^{3/2}} \sum_{a_1, a_3, a_4, a_5 \in \mathbf{V}(\underline{A})} \psi_1(-\tilde{h}_A(a_2, a_3, a_4))$$

Proof. By (39) we have that

(41)
$$\frac{1}{|k^*|} \sum_{\chi} \prod_{i=1}^{5} \tilde{F}(\theta\chi; \psi_i) = \frac{1}{|k^*|^6 |k|^{5/2}} \sum_{\chi_1,\ldots,\chi_5 \in \widehat{k^*}} \prod_{i=1}^{5} G(\chi_i, \chi_{i+1}; \psi_i)$$

Summing over $\chi_1, \ldots, \chi_5$ in (41), we get

$$(42) \qquad \frac{1}{|k^*||k|^{5/2}} \sum_{\substack{a_1,\ldots,a_5 \in V \\ x_1,\ldots,x_5 \in V}} \psi(\sum_{i=1}^{5} A_i a_i^{-1} x_i) =$$

$$\frac{1}{|k^*||k|^{5/2}} \sum_{\substack{a_1,\ldots,a_5 \in k^* \\ a_1+\cdots+a_5=0}} \sum_{x_1 \in k} \psi((\sum_{i=1}^{5} A_i a_i^{-1})x_1 + h(a_1,\ldots,a_5))$$

where

$$V = \left\{ \substack{a_1,\ldots,a_5 \in k^* \\ x_1,\ldots,x_5 \in k} : \substack{x_i - a_i = x_{i+1} + a_{i+1}, i=1,\ldots,4, \\ x_1 + a_1 = x_5 - a_5} \right\}$$

$$h(a_1,\ldots,a_5) = a_1 \left(\sum_{i=1}^{5} \frac{A_i}{a_i} \right) + 2 \sum_{j=2}^{4} a_j \left(\sum_{i=j+1}^{5} \frac{A_i}{a_i} \right)$$

Summing over x_1 gives
$$(43)$$
$$\frac{1}{|k^*||k|^{3/2}} \sum_{\substack{a_1,\ldots,a_5 \in k^* \\ a_1+\cdots+a_5=0 \\ \frac{A_1}{a_1}+\cdots+\frac{A_5}{a_5}=0}} \psi(h(a_1,\ldots,a_5)) = \frac{1}{|k|^{3/2}} \sum_{a_1,a_3,a_4,a_5 \in \mathbf{V}(\underline{A})} \psi(-\tilde{h}_A(a_2,a_3,a_4))$$

where the last equality is given by the change of variables $a_i \mapsto a_5^{-1} A_i a_i$, using the resulting equality

$$a_1^{-1} + a_2^{-1} + a_3^{-1} + a_4^{-1} + 1 = 0,$$

and summing over a_5 which no longer appears in the sum. $\qquad \square$

Proposition 6.4. *Let k be a finite field, and let ψ be an additive character of k. For $\underline{A} = (A_1, \ldots, A_5) \in k^*$, let $\mathbf{V}(A)$, $\tilde{h}_A$ as above. Denote by*

$$S_A(\psi) = \sum_{a_1,a_3,a_4,a_5 \in \mathbf{V}(\underline{A})} \psi_1(\tilde{h}_A(a_2, a_3, a_4))$$

Then there exist $M > 0$ such that

$$S_A(\psi) \le M|k|$$

Proof. In appendix A we prove the following properties of $\mathbf{V}(A), \tilde{h}_A$
   (1) For $A = (A_1, \ldots, A_5) \in (k^*)^5$ the variety $\mathbf{V}(A)$ is an irreducible two dimensional algebraic variety
   (2) Except 14 values, all fibers of $\tilde{h}_A : \mathbf{V}(A) \to \overline{k}$ are absolutely irreducible

and therefore the proposition is a corollary of theorem 3.3. $\qquad \square$

Theorem 4.3 is now a corollary of lemma 6.3 and proposition 6.4.

Remark 6.5. *Using the same averaging trick one can show that the third moment satisfies*

$$\frac{1}{p-1} \sum_{\chi} F(\chi; \psi_1) F(\chi; \psi_2) F(\chi; \psi_3) = p \sum_{a^2=B} \psi(a)$$

where $B = (A_1 + A_2 + A_3)^2 - 4(A_1 A_2 + A_1 A_3 + A_2 A_3)$

7. Matrix elements of the quantum cat map: Fluctuations in short windows

In this section we study the fluctuations of the matrix elements of quantum cat map about their limit. We prove the following theorem

Theorem 7.1. *Let* $A \in SL_2(\mathbb{Z})$ *be a hyperbolic matrix satisfying* $A \equiv I$ (mod 2). *Fix* $f \in C^\infty(\mathbb{T}^2)$ *of zero mean. Assume* $L < 2 \operatorname{ord}(A, N)$, *then as* $N \to \infty$ *through split primes satisfying* $\operatorname{ord}(A, N)/N^{2/3} \to \infty$, *the third moment of* $P(\theta)$ *satisfies*

$$(44) \qquad \int_0^1 \left(\sqrt{L} P(\theta) \right)^3 d\theta = O(\frac{N}{L^{3/2}})$$

We begin the proof by a reduction to the computation of mixed moments of $P_\nu(\theta)$. We show that it suffice to prove the following proposition

Proposition 7.2. *Let* $A \in SL_2(\mathbb{Z})$ *be a hyperbolic matrix satisfying* $A \equiv I$ (mod 2). *Fix* $0 \neq \nu_1, \nu_2, \nu_3 \in \mathbb{Z}$. *Then under the conditions of theorem 7.1*

$$(45) \qquad \int_0^1 (P_{\nu_1}(\theta) P_{\nu_2}(\theta) P_{\nu_3}(\theta))^3 d\theta = O(\frac{N}{L^3})$$

Theorem 7.1 is a consequence of this proposition as follows:
Let $f(x) = \sum_{n \in \mathbb{Z}^2} \hat{f}(n) e(nx) \in C^\infty(\mathbb{T}^2)$. Write $f(x) = f_N(x) + f_R(x)$ where

$$f_N(x) = \sum_{\|n\| < N^{1/4}} \hat{f}(n) e(nx), f_R(x) = \sum_{\|n\| \geq N^{1/4}} \hat{f}(n) e(nx)$$

Then we have that $P(\theta) = P_N(\theta) + P_R(\theta)$ correspondingly. By the fast decay of the Fourier coefficients of $f(x)$, $\|P_R(\theta)\|_\infty = O(\frac{1}{N^\infty})$, and by corollary 2.5 we have that

$$P_N(\theta) = \sum_{\nu \in \mathbb{Z}} f_N^\sharp(\nu) P_\nu(\theta) = \sum_{|\nu| < N} f^\sharp(\nu) P_\nu(\theta) + O(N^{-\infty})$$

again by the fast decay of the Fourier coefficients. By Cauchy-Schwartz inequality

$$\int_0^1 P^3(\theta) d\theta = \int_0^1 (P_N(\theta) + P_R(\theta))^3 d\theta = \int_0^1 P_N^3(\theta) d\theta + O_f(\frac{1}{N^\infty})$$

Now

$$\int_0^1 P_N^3(\theta) d\theta = \sum_{\nu_1, \nu_2, \nu_3 \in \mathbb{Z}} f^\sharp(\nu_1) f^\sharp(\nu_2) f^\sharp(\nu_3) \int_0^1 P_{\nu_1}(\theta) P_{\nu_2}(\theta) P_{\nu_3}(\theta) d\theta$$

which proves theorem 7.1 by proposition 7.2.

7.1. **Proof of proposition 7.2.** Denote

$$H_3(t_1, t_2) = \sum_{l_1, l_2 \in \mathbb{Z}} \widehat{h}(\frac{t_1 + l_1 \operatorname{ord}}{L}) \widehat{h}(\frac{t_2 + l_2 \operatorname{ord}}{L}) \widehat{h}(\frac{-t_1 - t_2 - (l_1 + l_2)\operatorname{ord}}{L})$$

Expanding the Fourier expansion of $P_\nu(\theta)$ and calculating the integral, we get that

(46)
$$\int_0^1 P_{\nu_1}(\theta) P_{\nu_2}(\theta) P_{\nu_3}(\theta) d\theta = \frac{1}{L^3} \sum_{\tau_1, \tau_2 \pmod{\operatorname{ord}(A,N)}} H_3(\tau_1, \tau_2) e(\bar{2} \frac{v(A^{\tau_1}, A^{\tau_2})}{N})$$

where

$$v(g_1, g_2) = q(k_1; g_1) + q(k_2; g_2) + q(k_3; g_1 g_2)$$

with $\omega(k_i, k_i A) = \nu_i, i = 1, 2, 3$. Since H_3 is periodic with period $\operatorname{ord}(A, N)$, we can write it as follows:

(47)
$$H_3(\tau_1, \tau_2) = \sum_{j_1, j_2 \pmod{\operatorname{ord}(A,N)}} \gamma(j_1, j_2) e(\frac{j_1 \tau_1 + j_2 \tau_2}{\operatorname{ord}(A, N)})$$

where

$$\gamma(j_1, j_2) = \frac{1}{\operatorname{ord}(A, N)^2} \sum_{\tau_1, \tau_2} \Gamma_3(\tau_1, \tau_2) e(\frac{-j_1 \tau_1 - j_2 \tau_2}{\operatorname{ord}(A, N)}) =$$

$$\frac{L^3}{\operatorname{ord}(A, N)^2} \int_0^1 h_L(x - \frac{j_1}{\operatorname{ord}(A, N)}) h_L(x - \frac{j_2}{\operatorname{ord}(A, N)}) h_L(x) dx$$

which are in particular positive. Plugging (47) in (46), and switching order of summation, we get that the RHS of (46) is

$$\frac{1}{L^{3/2}} \sum_{j_1, j_2 \pmod{\operatorname{ord}}(A,N)} \gamma(j_1, j_2) \sum_{\tau_1, \tau_2 \pmod{\operatorname{ord}}(A,N)} e(\frac{j_1 \tau_1 + j_2 \tau_2}{\operatorname{ord}(A, N)} + \frac{v(\lambda^{\tau_1}, \lambda^{\tau_2})}{N}) =$$

$$\frac{1}{L^{3/2}} \sum_{j_1, j_2 \pmod{\operatorname{ord}}(A,N)} \gamma(j_1, j_2) S_3(j_1, j_2)$$

where

(48)
$$S_3(j_1, j_2) = \sum_{\tau_1, \tau_2 \pmod{\operatorname{ord}(A,N)}} e(\frac{j_1 \tau_1 + j_2 \tau_2}{\operatorname{ord}(A, N)} + \frac{v(A^{\tau_i}, A^{\tau_2})}{N})$$

Proposition 7.2 will follow by showing that there exists $M > 0$ such that for any $j_1, j_2 \pmod{\operatorname{ord}}(A, N)$ $|S_3(j_1, j_2)| \le MN$. To show this we complete the sum to an exponential sum over the group $\mathcal{C}_A(N)$. For a pair of characters χ_1, χ_2 of $\mathcal{C}_A(N)$, and an additive character ψ of $\mathbb{F}_N$, set (as in (30))

(49)
$$\mathcal{S}(\chi_1, \chi_2; \psi) = \sum_{x, y \in \mathcal{C}_A(N)} {}' \chi_1(x) \chi_2(y) \psi(v(x, y))$$

Let ψ be the additive character satisfying $\psi(1) = e(\frac{\nu_1}{N})$ (recall $\nu_1 \neq 0$)), and let g be a generator of $\mathcal{C}_A(N)$ such that $g^r = A$, where $r = \frac{N}{\mathrm{ord}(A,N)}$. Denote by χ_0 the character of $\mathcal{C}_A(N)$ satisfying $\chi_0(g) = e(\frac{1}{N-1})$. We therefore get that $e(\frac{j\tau}{\mathrm{ord}(A,N)}) = \chi_0^j(A^\tau)$. Writing the indicator function of the subgroup generated by A as

$$\mathbb{1}_A(y) = \frac{N}{\mathrm{ord}(A,N)} \sum_{\substack{\theta \in \mathcal{C}_A(N) \\ \theta(A)=1}} \theta(y)$$

and $a = \nu_1^{-1}\nu_2, b = \nu_1^{-1}\nu_3$, we can write S_3 as

$$S_3 = \frac{N^2}{\mathrm{ord}(A,N)^2} \sum_{\substack{\theta_1,\theta_2 \in \mathcal{C}_A(N) \\ \theta_i(A)=1}} \mathcal{S}(\chi_1\theta_1, \chi_2\theta_2; \psi)$$

Proposition 7.3. *Let χ_1, χ_2 be characters of $\mathcal{C}_A(N)$, ψ an additive character of $\mathbb{F}_N$, and $v(x,y)$ as above. Then if A is diagonalizable over the finite field $\mathbb{F}_N$ then there exists $M > 0$ such that $|S_3(j_1, j_2)| \leq MN$.*

Proof. If A is diagonalizable over $\mathbb{F}_N$, then $\mathcal{C}_A(N) \simeq \mathbb{F}_N^*$. Under this isomorphism the sum $\mathcal{S}(\chi_1, \chi_2; \psi)$ becomes

$$\mathcal{S}(\chi_1, \chi_2; \psi) \sum_{x,y \in \mathbb{F}_N^*} \chi_1(x)\chi_2(y)\psi(\frac{1+x}{1-x} + a\frac{1+y}{1-y} + b\frac{xy+1}{xy-1})$$

By corollary 5.6 for the field $k = \mathbb{F}_N$, multiplicative characters $\chi_1^{j_1}\theta_1, \chi_1^{j_2}\theta_2$, and additive character ψ the claim follows $\qquad\square$

Corollary 7.4. *Let χ_1, χ_2 be characters of $\mathcal{C}_A(N)$, ψ an additive character of $\mathbb{F}_N$, and $v(x,y)$ as above. Then there exists $M > 0$ such that $|S_3(j_1, j_2)| \leq MN$.*

Proof. Consider the group $\mathcal{C}_A(N)$ as the $\mathbb{F}_N$ rational points of the algebraic group $\{B \in SL_2(\overline{\mathbb{F}_N}) : AB - BA = 0\}$. Denote by $w_1, \ldots, w_l$ the roots of the exponential sum $\mathcal{S}(\chi_1, \chi_2; \psi)$. Denote by σ the Frobenius automorphism of $\mathbb{F}_{N^2}/\mathbb{F}_N$, and for $x \in \mathcal{C}_A(N^2)$ denote by $\mathcal{N}(x) = x\sigma(x)$. Then by Deligne's result we have that

$$\mathcal{S}_2(\chi_1, \chi_2; \psi) = \sum_{x,y \in \mathcal{C}_A(N^2)} \chi_1 \circ \mathcal{N}(x)\chi_2 \circ \mathcal{N}(y)\psi(\mathrm{tr}(v(x,y))$$

satisfies

$$\mathcal{S}_2 = w_1^2 + \cdots + w_l^2$$

Over $\mathbb{F}_{N^2}$ the matrix A is diagonalizable, and therefore we can apply proposition 7.3, and get that $|w_i^2| \leq N^2, i = 1, \ldots, l$, hence $|w_i| \leq N$, and the corollary follows. $\qquad\square$

We now have that

$$(50) \qquad \int_0^1 (P(\theta))^3 \, d\theta = O\left(\frac{N}{L^3} \sum_{j_1, j_2} |\gamma(j_1, j_2)| \right)$$

and since $\gamma(j_1, j_2)$ are positive we can drop the absolute value and remain with $\sum |\gamma(j_1, j_2)| = \Gamma(0) = O(1)$ since $L < 2 \operatorname{ord}(A, N)$. This concludes the proof.

8. Hecke matrix elements: Independence

In [19] Kurlberg and Rudnick showed that for any hyperbolic matrix $A \in SL_2(\mathbb{Z})$, and for any N there exist a basis $\{\psi_i\}_{i=1}^N$ of $U_N(A)$ called (the) Hecke basis satisfying that for any smooth function $f \in C^\infty(\mathbb{T}^2)$

$$(51) \qquad |\langle \operatorname{Op}_N(f)\psi_i, \psi_i \rangle - \int_{\mathbb{T}^2} f| \ll N^{-1/4 - \epsilon}$$

This result was later improved by Hadani and Gurevich for N prime to

$$(52) \qquad |\langle \operatorname{Op}_N(f)\psi_i, \psi_i \rangle - \int_{\mathbb{T}^2} f| \ll N^{-1/2}$$

It was later conjectured by them ([22]) that when normalizing these matrix elements by the correct size of N, and add together the Fourier coefficients that correspond to natural symmetries of the system, the fluctuations become equidistributed and independent in the semiclassical limit. In this section we prove some agreement with this conjecture. To state the precise theorem we start with a some background that was not covered in section 2. As a general reference we use [19, 20, 22]

8.1. **Hecke Theory for the Quantum cat map.** For a hyperbolic matrix $A \in SL_2(\mathbb{Z})$ satisfying $A \equiv I \pmod 2$, and N prime Let
$$(53)$$
$$\mathcal{C}_A(N) = SO(Q, \mathbb{Z}/N\mathbb{Z}) = \{B \in SL_2(\mathbb{Z}/N\mathbb{Z}) : AB = BA \pmod N\}$$

where $Q(n) = \omega(n, nA)$. The set

$$\{U_N(B) : B \in \mathcal{C}_A(N)\}$$

Is called the Hecke operators, and a basis of joint eigenfunction of all Hecke operators is called a Hecke basis.

Lemma 8.1. *Let $\{\psi_j\}_{j=1}^N$ be a Hecke basis, and let $m, n \in \mathbb{Z}^2$ such that $Q(n) = Q(m)$, then for all sufficiently large primes N we have*

$$(-1)^{n_1 n_2} \langle T_N(n)\psi_j, \psi_j \rangle = (-1)^{m_1 m_2} \langle T_N(m)\psi_j, \psi_j \rangle, \quad j = 1, \ldots, N$$

In light of this lemma, for $\nu \in \mathbb{Z}$, and ψ a Hecke eigenfunction define

$$Y_\nu(\psi) = \sqrt{N}(-1)^{n_1 n_2} \langle T_N(n)\psi, \psi \rangle$$

where $n \in \mathbb{Z}^2$ is such that $Q(n) = \nu$ if it exists (This is well defined by lemma 8.1). With this definition conjecture 2.1 is the same as

Conjecture 8.2. *As $N \to \infty$ through primes, ,for any $\nu \in Z$, the normalized matrix element $Y_\nu(\psi)$ has a limiting distribution of $\mathrm{tr}(U_\nu)$ as in conjecture 2.1. Moreover, the sequence*

$$\ldots, Y_{-3}(\psi),\ Y_{-2}(\psi),\ Y_{-1}(\psi), Y_1(\psi),\ Y_2(\psi),\ Y_3(\psi), \ldots$$

converge to a sequence of IID random variables

In this section we prove the following theorem, which is in agreement with conjecture 2.1.

Theorem 8.3. *Let $0 \neq \nu_1, \ldots, \nu_5 \in \mathbb{Z}$. Let N be a prime number, let $\{\psi_j\}_{j=1}^N$ be a Hecke basis of $U_N(A)$, then*

(54)
$$\frac{1}{N} \sum_{j=1}^N Y_{\nu_1}(\psi_j) \ldots Y_{\nu_5}(\psi_j) \ll \frac{1}{\sqrt{N}}$$

Remark 8.4. *We note that theorem 2 is a consequence of this theorem, by similar arguments to those showing that theorem 7.1 is a consequence of proposition 7.2.*

8.2. **Averaging operator.** For $n \in \mathbb{Z}^2$ let

$$D(n) = \frac{1}{\mathcal{C}_A(N)} \sum_{B \in \mathcal{C}_A(N)} T_N(nB)$$

The following lemma shows that this averaging operator is essentially diagonal with respect to a Hecke basis (for proof see lemma 7 in [22]).

Lemma 8.5. *Let $\tilde{D}$ be the matrix obtained when expressing $D(n)$ in terms of the Hecke eigenbasis. Then $\tilde{D}$ has the form*

$$\tilde{D} = \begin{pmatrix} D_{11} & D_{12} & 0 & 0 & \ldots & 0 \\ D_{21} & D_{22} & 0 & 0 & \ldots & 0 \\ 0 & 0 & D_{33} & 0 & \ldots & 0 \\ 0 & 0 & 0 & D_{44} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & D_{NN} \end{pmatrix}$$

such that

$$|D_{ij}| \ll N^{-1/2}$$

for $1 \leq i, j \leq 2$.

Lemma 8.6. *Let $\{\psi_j\}_{j=1}^N$ be a Hecke basis of $\mathcal{H}_N$, and let $0 \neq n_1, \ldots, n_5 \in \mathbb{Z}^2$. Then*

$$\sum_{j=1}^N \langle T_N(n_1)\psi_j, \psi_j \rangle \cdots \langle T_N(n_1)\psi_j, \psi_j \rangle = \mathrm{tr}(D(n_1) \cdots D(n_5)) + O(N^{-5/2})$$

Proof. By lemma 8.5, we have that

$$\mathrm{tr}(D(n_1)\cdots D(n_5)) = \sum_{j=3}^{N} D(n_1)_{jj}\cdots D(n_5)_{jj} + \mathrm{tr}(A_1\ldots A_5)$$

where $A_1,\ldots,A_5$ are 2×2 matrices defined by

$$(A_k)_{ij} = (D(n_k))_{ij} \quad, 1\le i,j\le 2.$$

By lemma 8.5

$$\mathrm{tr}(A_1\ldots A_5) = D(n_1)_{11}\cdots D(n_5)_{11} + D(n_1)_{22}\cdots D(n_5)_{22} + O(N^{-5/2})$$

and by definition of $D(n)$ the proof is concluded. $\qquad\square$

8.3. **Proof of theorem 8.3.** The first step of the proof is to reduce the required moment into an exponential sum.

Lemma 8.7. *Choose* $n_1,\ldots,n_5$ *such that* $Q(n_i) = \nu_i$. *Then For* $N > N_0(\nu_1,\ldots,\nu_5)$ *we have*
(55)
$$\frac{1}{N}\sum_{j=1}^{N} V_{\nu_1}(\psi_j)\cdots V_{\nu_5}(\psi_j) = \frac{N^{5/2}}{|\mathcal{C}_A(N)|^5}\sum_{\substack{B_1,\ldots,B_5\in\mathcal{C}_A(N)\\ n_1 B_1+\cdots+n_5 B_5=0 \pmod N}} e(\frac{\bar{2}u(B_1,\ldots,B_5)}{N})+O(N^{-1})$$

where $\bar{2}$ *is the inverse of* 2 *mod* N, *and*

$$u(B_1,B_2,B_3,B_4,B_5) = \sum_{1\le i<j\le 5}\omega(n_i B_i, n_j B_j)$$

Proof. By definition we have that

$$\frac{1}{N}\sum_{j=1}^{N} V_{\nu_1}(\psi_j)\cdots V_{\nu_5}(\psi_j) =$$

$$N^{3/2}\epsilon(n_1)\ldots\epsilon(n_5)\sum_{j=1}^{N}\langle T_N(n_1)\psi_j,\psi_j\rangle\cdots\langle T_N(n_5)\psi_j,\psi_j\rangle =$$

$$N^{3/2}\epsilon(n_1)\ldots\epsilon(n_5)\,\mathrm{tr}(D(n_1)\cdots D(n_5)) + O(N^{-1})$$

where the last equality is by lemma 8.6. By definition of $D(n)$, and by (3),(4) we have

$$\mathrm{tr}(D(n_1)\cdots D(n_5)) = \frac{1}{|\mathcal{C}_A(N)|^5}\sum_{B_1,\ldots,B_5\in\mathcal{C}_A(N)} T_N(n_1 B_1)\cdots T_N(n_5 B_5) =$$

$$\frac{N}{|\mathcal{C}_A(N)|^5}\epsilon(n_1 B_1)\ldots\epsilon(n_5 B_5)\sum_{\substack{B_1,\ldots,B_5\in\mathcal{C}_A(N)\\ n_1 B_1+\cdots+n_5 B_5=0 \pmod N}} e(\frac{\bar{2}u(B_1,\ldots,B_5)}{N})$$

and since if $B\in\mathcal{C}_A(N)$ then $B\equiv I \pmod 2$, we have that $\epsilon(n_i) = \epsilon(n_i B_i)$, which concludes the proof. $\qquad\square$

Theorem 8.3 is now a consequence of the following proposition

Proposition 8.8. *If $Q(n_i) \neq 0, i = 1, \ldots, 5$, then for sufficiently large split prime N*

$$(56) \qquad \sum_{\substack{B_1,\ldots,B_5 \in \mathcal{C}_A(N) \\ n_1 B_1 + \cdots + n_5 B_5 = 0 \pmod N}} e(\frac{\bar{2}u(B_1,\ldots,B_5)}{N}) \ll N^2$$

Proof. By assumption there exist $M \in SL_2(\mathbb{F}_N)$ such that $MAM^{-1} = D$ where D is diagonal. In this case the group $\mathcal{C}_A(N)$ is conjugated to the group

$$\mathcal{C}_A(N)^M = \{MBM^{-1} : B \in \mathcal{C}_A(N)\} = \{\begin{pmatrix} x & \\ & x^{-1} \end{pmatrix} : x \in \mathbb{F}_N^*\}$$

we can therefore write the sum as

$$(57) \qquad \sum_{\substack{D_1,\ldots,D_5 \in \mathcal{C}_A(N)^M \\ n_1 D_1 + \cdots + n_5 D_5 = 0 \pmod N}} e(\frac{\bar{2}u(MB_1M^{-1},\ldots,MB_5M^{-1})}{N})$$

Writing $D_i = \begin{pmatrix} x_i & \\ & x_i^{-1} \end{pmatrix}$ and $m^{(i)} = (m_1^{(i)}, m_2^{(i)}) := n_i M$, we have that

$$u(MB_1M^{-1},\ldots,MB_5M^{-1}) = \sum_{1 \leq i < j \leq 5} \omega(n_i MD_i, n_j MD_j) =$$

$$\sum_{q \leq i < j \leq 5} \left(m_1^{(i)} m_2^{(j)} x_i x_j^{-1} - m_2^{(i)} m_1^{(j)} x_j x_i^{-1} \right)$$

and the condition $n_1 B_1 + \cdots + n_5 B_5 = 0 \pmod N$ becomes

$$\left\{ \begin{matrix} m_1^{(1)} x_1 + \cdots + m_1^{(5)} x_5 = 0 \\ m_2^{(1)} x_1^{-1} + \cdots + m_2^{(5)} x_5^{-1} = 0 \end{matrix} \right\} = \left\{ \begin{matrix} m_1^{(1)} m_2^{(1)} x_1 + \cdots + m_1^{(5)} m_2^{(5)} x_5 = 0 \\ x_1^{-1} + \cdots + x_5^{-1} = 0 \end{matrix} \right\}$$

where the equations are in $\mathbb{F}_N$, and the last equality is by the change of variables $x_i \mapsto m_2^{(i)} x_i$. Denote $A_i = m_1^{(i)} m_2^{(i)}$, (notice that these are nonzero since n_i is not an eigenvector of A, as $Q(n_i) \neq 0$) (57) is now

$$(58) \qquad \sum_{\substack{0 \neq x_1,\ldots,x_5 \in \mathbb{F}_N \\ A_1 x_1 + \cdots + A_5 x_5 = 0 \\ x_1^{-1} + \cdots + x_5^{-1} = 0}} e(\frac{\tilde{h}_A(x_3, x_4, x_5)}{N})$$

where

$$\tilde{h}_A(x_3, x_4, x_5) = \frac{1}{x_3} + \frac{1}{x_4} + \frac{1}{x_5} + A_3 x_3 \left(\frac{1}{x_4} + \frac{1}{x_5} \right) + \frac{A_4 x_4}{x_5}$$

as in lemma 6.3. By proposition 6.4 the result now follows. $\qquad \square$

Corollary 8.9. *Let $n_1, \ldots, n_5 \in \mathbb{Z}^2$ such that $Q(n_i) \neq 0, i = 1, \ldots, 5$, then for all primes N not dividing $\mathrm{tr}(A)^2 - 4$ the inequality in (56) holds*

Proof. By proposition 8.8 we have that if A is diagonalizable over k, then

$$S(A) := \sum_{\substack{B_1,\ldots,B_5 \in \mathcal{C}_A(N) \\ n_1 B_1 + \cdots + n_5 B_5 = 0 \pmod{N}}} e(\frac{\bar{2}u(B_1,\ldots,B_5)}{N}) \ll |k|^2$$

and in particular all roots of the corresponding L-function $L(S(A))$ are of absolute value at most $|k|^2$. We now show that the diagonalizable condition may be dropped using "base change". For a field k consider the algebraic variety over the algebraic closure $\overline{k}$ of k defined by

$$\mathcal{C}_A(\overline{k}) = \{B \in SL_2(\overline{k}) : BA = AB\}$$

and for $n_1,\ldots,n_5 \in k^2$ the subvariety of $\mathcal{C}_A(\overline{k})^5$ defined by

$$V = \{(B_1,\ldots,B_5) \in \mathcal{C}_A(\overline{k})^5 : n_1 B_1 + \cdots + n_5 B_5 = 0\}$$

For any finite extension of k of degree ν, k_ν, let

$$V(k_\nu) = \{(B_1,\ldots,B_5) \in \mathcal{C}_A(k_\nu) : (B_1,\ldots,B_5) \in V\}$$

be the k_ν-rational points of V. Let $\omega_1,\ldots,\omega_l$ be the roots of $L(S(A))$. Then by Deligne's result we have that

$$\sum_{(B_1,\ldots,B_5) \in V(k_\nu)} e(\frac{\mathrm{tr}_{k_\nu/k}(\bar{2}u(B_1,\ldots,B_5))}{N}) = \omega_1^\nu + \cdots + \omega_l^\nu$$

If $\mathrm{tr}(A)^2 \neq 4$, then for $\nu = 2$, A is diagonalizable over k_ν, and hence by proposition 8.8 $|\omega_i^2| \leq |k|^4, i = 1,\ldots,l$, and therefore $|\omega_i| \leq |k|^2$ which concludes the proof of proposition 8.8 for all primes N that do not divide $\mathrm{tr}(A)^2 - 4$. □

9. DISCUSSION

9.1. **Matrix elements and exponential sums.** The connection between the matrix elements of the cat map and the family of exponential sum $F(\chi;\psi)$, was observed previously by Kurlberg and Rudnick, Gurevich and Hadani, and Kelmer ([22, 9, 15]). In [15], Kelmer shows that the Hecke matrix element corresponding to the (non quadratic) character χ of $\mathcal{C}_A(N)$ is in fact of the form of $F(\chi;\psi)$, hence at least in the split case they coincide. Therefore conjecture 2.1 can be interpreted in the split case as prediction to the value distribution of this family as χ varies. Conjecture 4.2 is a generalization of this conjecture and, predicts that the action of the group of characters on this family has a '*mixing type*' behaviour (conjecture 4.2.2). Agreement with these predictions can be seen in figure 3. It shows high agreement of the numerical plots of mixed sixth moments, and fourth moment. In the left part moments of type

$$\frac{1}{p-1} \sum_{\chi} F(\chi_1\chi;\psi) \cdots F(\chi_6\chi;\psi)$$

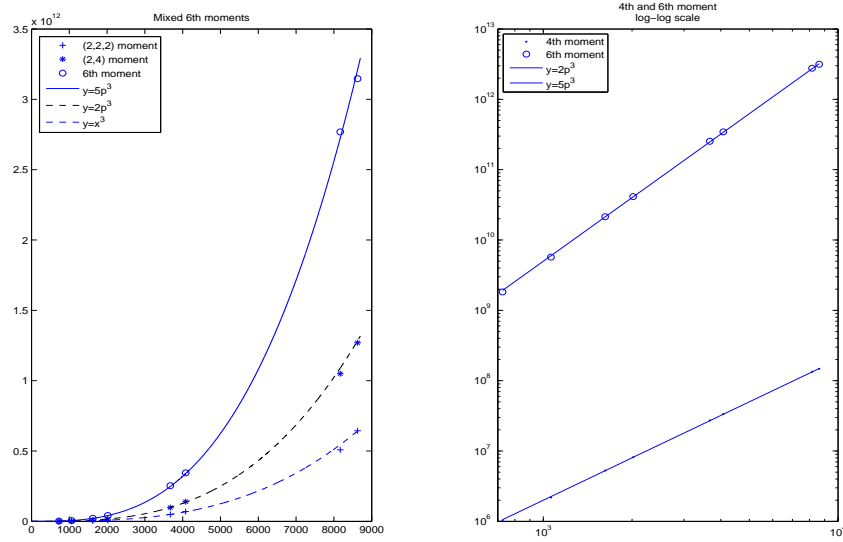FIGURE 3. High mixed moments

in different pairing. In case all characters $\chi_1, \ldots, \chi_6$ are equal, it shows asymptotic growth of $5p^3$, and in case the characters are split into subset of two and four equal characters (3 different pairs) we see asymptotic growth of $2p^3$ (respectively p^3). This shows agreement with conjecture 4.2 recalling that if X is a random variable with Sato Tate distribution, then

$$\mathbb{E}(X^{2n}) = \begin{cases} 1 & n=1 \\ 2 & n=2 \\ 5 & n=3 \end{cases}$$

It is a generic assumptions on matrix elements that they behave independently with respect to the eigenfunctions. When translating conjecture 4.2 to the language of Hecke matrix elements, this independence behaviour appears as follows. In the case where N is prime, one can define a "product law" for the Hecke eigenfunctions, by parameterizing the Hecke eigenfunction using the characters of $\mathcal{C}_A(N)$, ψ_χ. We define for every character χ_1 of $\mathcal{C}_A(N)$ the following operator

$$\begin{aligned} M_\chi : \mathcal{H}_N &\rightarrow \mathcal{H}_N \\ \psi_\chi &\mapsto \psi_{\chi_1 \chi} \end{aligned}$$

Conjecture 4.2 is therefore: for fixed $l \in \mathbb{N}$ and for any prime N choose l characters $\chi_1, \ldots, \chi_l$ of $\mathcal{C}_A(N)$. Then as $N \to \infty$ through primes

$$\frac{1}{\mathcal{C}_A(N)} \sum_{\chi \in \mathcal{C}_A(N)} Y_{\nu_1}(L_{\chi_1}(\psi_\chi)) \cdots Y_{\nu_l}(L_{\chi_l}(\psi_\chi)) \to \mathbb{E}(\mathrm{tr}(U_{\nu_1}) \cdots \mathrm{tr}(U_{\nu_l}))$$

where U_ν are as in conjecture 2.1 independent for $\nu_i \neq \nu_j$.

9.2. **Fluctuations in short windows.** The expected independence behaviour of the matrix elements suggests more on the fluctuations in short windows. Since at every point we sum matrix elements related to different eigenvalues (and hence they have independent behaviour), a Gaussian limiting distribution may appear. The following figures show agreement with this heuristic.   In figure 4 comparison between the distribution of $P(\theta)$,
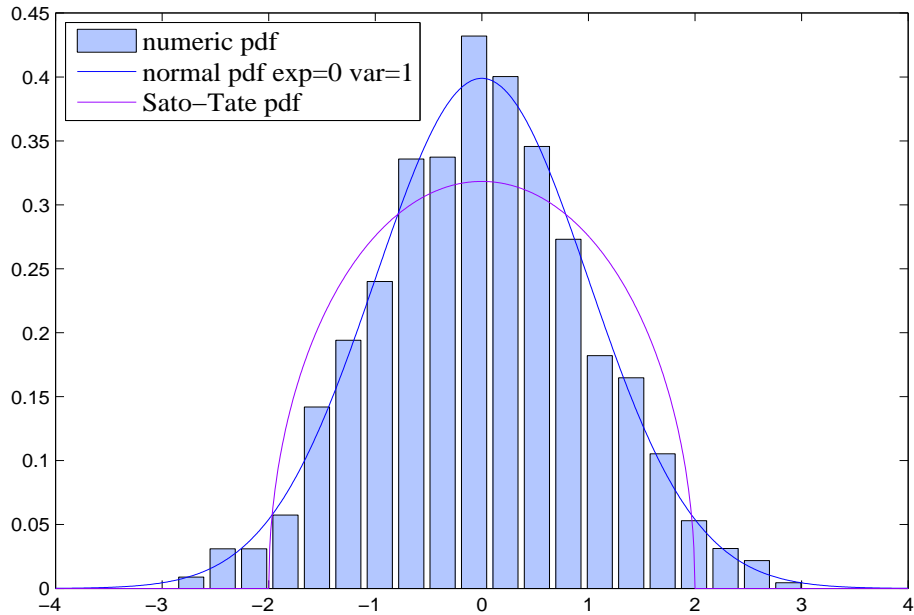

FIGURE 4. $P(\theta)$ distribution, $f(x) = e(x + y)$

normal distribution, and Sato-Tate distribution is displayed. It shows that the distribution agrees with normal distribution rather than Sato-Tate. In figure 5 the function $f(x)$ is a trigonometrical polynomial with exponents that give two different values for $Q(n)$. We thus see that the variance is 2 rather than 1 as in figure 4. This is in fact the result shown in [18]. In fact, assuming conjecture 4.2, it is possible to show the following theorem (see [27]):

Theorem 9.1. *Let $A \in SL_2(\mathbb{Z})$ be a unimodular matrix with distinct eigenvalues such that $A \equiv I \pmod 2$. Fix a smooth function $f \in C^\infty(\mathbb{T}^2)$. Assume that for every $\epsilon > 0$ $N^{1-\epsilon} \ll L$, and $\frac{L}{\mathrm{ord}(A,N)} \to 0$ as N goes to*
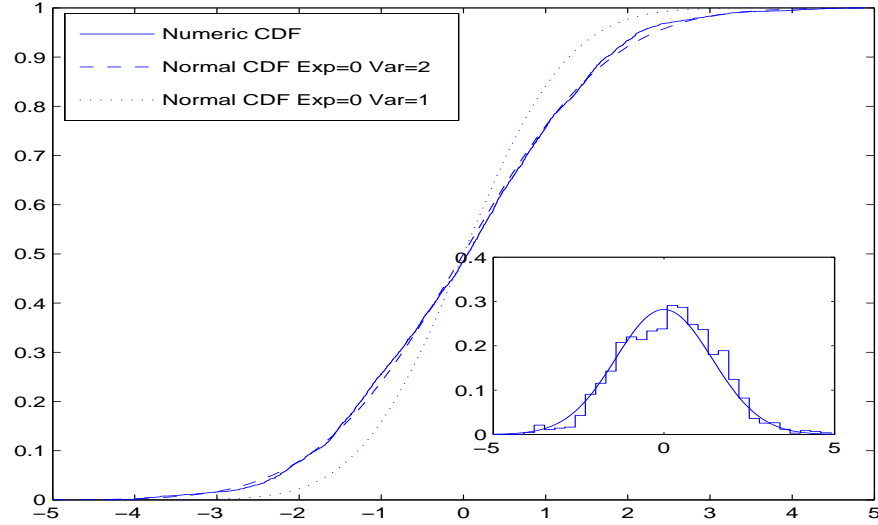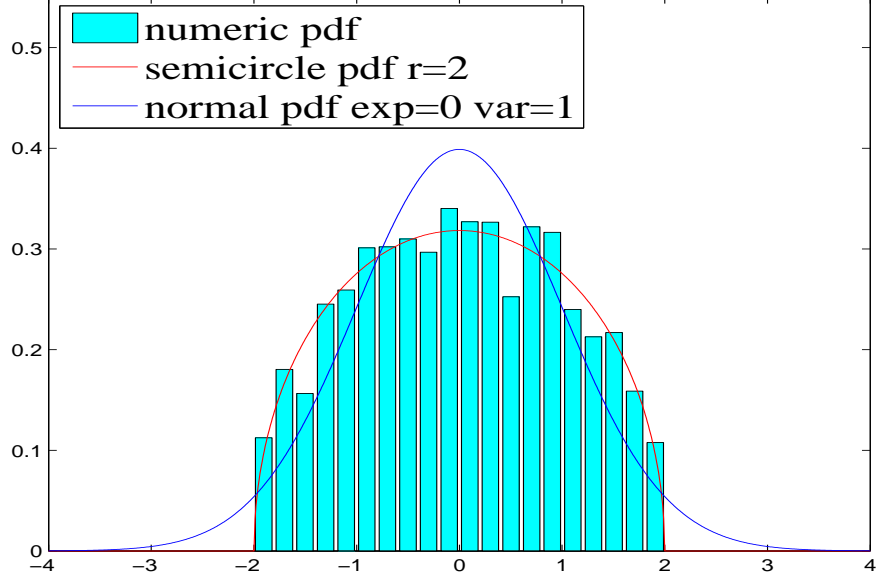
FIGURE 5. $P(\theta)$ distribution, $f(x) = e(x + y) + e(2x + y)$

infinity through primes, then, assuming conjecture 4.2, $P(\theta)$ has Gaussian limiting distribution with mean 0 and $var(P) = \sum_{\nu \in \mathbb{Z}} f^{\sharp}(\nu)$

Remark 9.2. *Notice that the conditions $L/\operatorname{ord}(A, N) \to 0$ and $N^{1-\varepsilon} \ll L$ imply an assumption on the size of $\operatorname{ord}(A, N)$, however, assuming GRH this assumption is valid for most primes (c.f.[17]).*

This result is in some agreement with predicted results on generic systems. It says that once the arithmetic symmetries of the systems are grouped together, the resulting desymetrized components have a generic Gaussian limiting behaviour. However we should notice that when the size of the window becomes too short, the function $P(\theta)$ no longer consists of sums of matrix elements corresponding to different eigenvalues, and in cases where the order of A modulo N is maximal it studies the matrix elements distribution and we no longer expect normal distribution but rather Sato-Tate as is shown in figure 6

FIGURE 6. $P(\theta)$ distribution, Very short window

APPENDIX A. PROOFS OF IRREDUCIBILITY

Let k be a finite field, and $0 \neq A_1, \dots, A_5 \in k$. Denote by

$$\mathbf{V}(A) = \left\{ 0 \neq a_1, a_2, a_3, a_4 \in k : \begin{smallmatrix} A_1a_1 + A_2a_2 + A_3a_3 + A_4a_4 + A_5 = 0 \\ a_1^{-1} + a_2^{-1} + a_3^{-1} + a_4^{-1} + 1 = 0 \end{smallmatrix} \right\}$$

and by

$$\mathbf{V}(A, \nu) = \left\{ 0 \neq a_1, a_2, a_3, a_4 \in k_\nu : \begin{smallmatrix} A_1a_1 + A_2a_2 + A_3a_3 + A_4a_4 + A_5 = 0 \\ a_1^{-1} + a_2^{-1} + a_3^{-1} + a_4^{-1} + 1 = 0 \end{smallmatrix} \right\}$$

$$\overline{\mathbf{V}}(A) = \left\{ 0 \neq a_1, a_2, a_3, a_4 \in \overline{k} : \begin{smallmatrix} A_1a_1 + A_2a_2 + A_3a_3 + A_4a_4 + A_5 = 0 \\ a_1^{-1} + a_2^{-1} + a_3^{-1} + a_4^{-1} + 1 = 0 \end{smallmatrix} \right\}$$

the rational points in any finite extension k_ν of k, and the points in the algebraic closure $\overline{k}$. Let

$$\tilde{h}_A(a_2, a_3, a_4) = \sum_{2 \leq i < j \leq 4} \left(\frac{A_i a_i}{a_j} - \frac{A_j a_j}{a_i} \right) + \sum_{i=1}^{4} A_i a_i - A_5 \sum_{i=2}^{4} a_i^{-1}$$

We prove the following lemmas

Lemma A.1. *For $0 \neq A_1, \dots, A_5 \in k$, the variety $\overline{\mathbf{V}}(A)$ has one irreducible component of dimension 2*

and

Lemma A.2. *Except for 14 values of $C \in \overline{k}$ the fibers*

$$\tilde{h}_A^{-1}(C) \subset \overline{\mathbf{V}}(A)$$

are irreducible

A.1. **Proof of lemma A.1.** We prove the lemma by counting the number of points on $\mathbf{V}(A, \nu)$ for every $\nu \in \mathbb{N}$. By Lang-Weil theorem, any irreducible component surface has $|k_\nu|^2 + O(|k_\nu|^{3/2})$ points on it, and therefore by showing this we prove the lemma. For a nontrivial additive character ψ of k_ν, and $0 \neq a, b \in k_\nu$ denote by $Kl(a, b)$ the Kloosterman sum

$$Kl(a, b) = \sum_{0 \neq x \in k_\nu} \psi(ax + bx^{-1})$$

Proposition A.3. *Let ψ be a nontrivial additive character of k_ν. Then*

$$\sharp \mathbf{V}(A) - |k_\nu|^2 = \frac{1}{|k_\nu|^2} \sum_{b \in k_\nu} \left(\prod_{i=1}^{5} Kl(a_i, b) \right) + O(|k_\nu|^{-1})$$

where $Kl(a, b) = \sum_{0 \neq x} \psi(ax + bx^{-1})$, is the Kloosterman sum.

Proof. By the orthogonality relations of additive characters, we have that

$$\sharp \mathbf{V}(A) = \frac{1}{|k_\nu|^2} \sum_{a, b \in k_\nu} \sum_{\underline{x} \in (k_\nu^*)^4} \psi(af_1(\underline{x}) + bf_2(\underline{x})) =$$

$$|k_\nu|^2 + \frac{1}{|k_\nu|^2} \sum_{b \in k_\nu^*} \psi(-b) \prod_{i=1}^{4} Kl(0, b) +$$

$$\frac{1}{|k_\nu|^2} \sum_{a, b \in k_\nu, a \neq 0} \psi(aa_5 + b) \prod_{i=2}^{5} (Kl(aa_i, b)) + O(|k_\nu|^{-1})$$

Using that $Kl(0, b) = -1, Kl(ac, b) = Kl(c, ab)$ we have that

$$\sharp \mathbf{V}(A) - |k_\nu|^2 = \frac{1}{|k_\nu|^2} \sum_{a \in k_\nu^*} \sum_{b \in k_\nu} \psi(-aa_5 - b) \prod_{i=1}^{4} (Kl(a_i, ba))^4 + O(|k_\nu|^{-1})$$

and under the change of variable $b \mapsto b/a$, we get

$$\sharp \mathbf{V}(A) - |k_\nu|^2 = \frac{1}{|k_\nu|^2} \sum_{b \in k_\nu} \prod_{i=1}^{5} (Kl(a_i, b)) + O(|k_\nu|^{-1})$$

and now using Weil's bound $Kl(a_i, b) \leq 2N^{\nu/2}$, we get the bound

$$\sharp \mathbf{V}(A) = |k_\nu|^2 + O(|k_\nu|^{3/2})$$

which proves the lemma. □

A.2. **Proof of lemma A.2.** We prove the irreducibility of the fibers by the following strategy: For each curve $\tilde{h}_A^{-1}(C)$ we find a curve in the affine plane $\mathbb{A}$ over $\overline{k}$ given by the zeros set of a polynomial, such that (an open Zariski subset of) the fiber is parameterized by (an open Zariski subset of) this plane curve. We then show that the polynomial defining the plane curve is irreducible over $\overline{k}$ and thus proving the lemma.

For simplicity of notations we use the following notation: For a polynomial $P(X_1, \ldots, X_n) \in k[X_1, \ldots, X_n]$ over a field k we denote the zeros set of this polynomial by

$$Z(p) = \{(a_1, \ldots, a_n \in \mathbb{A}^n : P(a_1, \ldots, a_n) = 0\}$$

and its complement by

$$Y_p = \{(a_1, \ldots, a_n) \in \mathbb{A}^n : P(a_1, \ldots, a_n) \neq 0\}$$

For fixed $A_1, \ldots, A_5, C \in k$ we define the following polynomial $p(a_3, a_4)$:

$$(59) \quad p(a_3, a_4) = 2a_4^2((B + C - 2A_2)a_3 + 2A_4a_4)(A_3a_3 + A_4a_4) +$$
$$a_4(2(B - C - 2A_2 + 2A_5)(A_4a_4^2 + A_3a_3^2) + (B^2 - C^2 + D)a_3a_4) +$$
$$2A_5((B - C + 2A_2)a_4 + 2A_3a_3)(a_3 + a_4)$$

where $B = A_2 + A_3 + A_4 + A_5 - A_1, D = -4A_2(A_3 + A_4 + A_5) + 4(A_3A_4 + A_3A_5 + A_4A_5)$.

Proposition A.4. *Let $p(a_3, a_4)$ be as above. Define*

$$\begin{aligned} g_1(x_3, x_4) &= A_3x_3 + A_4x_4 + A_5 \\ g_2(x_3, x_4) &= (B + C)x_3x_4 + 2A_4x_4^2 + 2A_5(x_4 + x_3) \end{aligned}$$

and

$$\begin{aligned} \tilde{g}_1(a_1, a_2, a_3, a_4) &= A_3a_3 + A_4a_4 + A_5 \\ \tilde{g}_2(a_1, a_2, a_3, a_4) &= (B + C)a_3a_4 + 2A_4a_4^2 + 2A_5(a_4 + a_3) \end{aligned}$$

Then the following map

$$a_1 = -\frac{A_2a_2(x_3, x_4) + A_3x_3 + A_4x_4 + A_5}{A_1}$$
$$a_2 = \frac{-2x_3x_4g_1(x_3, x_4)}{g_2(x_3, x_4)}$$
$$a_3 = x_3$$
$$a_4 = x_4$$

defines a bijection between $\tilde{h}_A^{-1}(C) \cap Y_{\tilde{g}_1} \cap Y_{\tilde{g}_2}$ and $Z(p) \cap Y_{g_1} \cap Y_{g_2}$

Proof. It is straightforward to check that if $(x_3, x_4) \in Z(p)$, then their lies in $\tilde{h}_A^{-1}(C)$. To show the other direction, we consider the system of equations

$$A_1 a_1 + A_2 a_2 + A_3 a_3 + A_4 a_4 + A_5 = 0$$

$$\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \frac{1}{a_4} + 1 = 0$$

$$\sum_{2 \le i < j \le 4} \left(\frac{A_i a_i}{a_j} - \frac{A_j a_j}{a_i} \right) + \sum_{i=1}^{4} A_i a_i - A_5 \sum_{i=2}^{4} a_i^{-1} = C$$

Multiply the second equation by $A_1 a_1 a_2 a_3 a_4$ and substitute $A_1 a_1$ by $-(A_2 a_2 - A_3 a_3 - A_4 a_4 - A_5)$ to get

$$(60) \quad A_2 a_2^2 (a_3 + a_4 + a_3 a_4) + A_3 a_3^2 (a_2 + a_4 + a_2 a_4) +$$
$$A_4 a_4^2 (a_2 + a_3 + a_2 a_3) + B a_2 a_3 a_4 + A_5 (a_2 a_3 + a_2 a_4 + a_3 a_4) = 0.$$

Next we multiply the third equation by $a_2 a_3 a_4$ to get

$$(61) \quad A_2 a_2^2 (a_3 + a_4 + a_3 a_4) + A_3 a_3^2 (-a_2 + a_4 + a_2 a_4) +$$
$$A_4 a_4^2 (-a_2 - a_3 + a_2 a_3) - A_5 (a_2 a_3 + a_2 a_4 + a_3 a_4) = 0.$$

Subtract (60) by (61) to get:

$$g_2(a_3, a_4) a_2 + 2 a_4 a_3 g_1(a_3, a_4) = 0$$

By assumption that $g_2(a_3, a_4) \neq 0$ we get that

$$a_2 = \frac{-2 a_3 a_4 g_1(a_3, a_4)}{g_2(a_3, a_4)}$$

Use this expression for a_2 inside $\tilde{h}_A(a_2, a_3, a_4) - C = 0$ to get

$$\frac{p(a_3, a_4)}{2 a_4 g_2(a_3, a_4)} = 0$$

and therefore a_3, a_4 must satisfy $p(a_3, a_4) = 0$. $\qquad \square$

Proposition A.5. *Let $p(X_3, X_4)$ as in (59). Then for all $C \in \overline{k}$ but at most 14 values, the polynomial $p(X_3, X_4)$ is irreducible over $\overline{k}$.*

Proof. We first denote the following homogeneous parts of $p(X_3, X_4)$

$$p_4(X_3, X_4) = 2X_4^2((B + C - 2a_2)X_3 + 2a_4 X_4)(a_3 X_3 + a_4 X_4)$$
$$p_3(X_3, X_4) = X_4(2(B - C - 2(a_2 - a_5)))(a_4 X_4^2 + a_3 X_3^2) +$$
$$(B^2 - C^2 + D)X_3 X_4)$$
$$p_2(X_3, X_4) = 2a_5((B - C + 2a_2)X_4 + 2a_3 X_3)(X_3 + X_4)$$

Let $q(X_3, X_4), r(X_3, X_4)$ be two polynomials satisfying $p(X_3, X_4) = q(X_3, X_4)r(X_3, X_4)$, and denote their decomposition into homogeneous parts, $q = q_0 + q_1 + q_2 + q_3, r = r_0 + r_1 + r_2 + r_3$. Without loss of generality we assume $\deg r \le \deg(q)$. Since the homogeneous parts of p are of degree 2,3,4 only this imposes a few restrictions on q, r. We split the cases into 2 parts

(1) **Case 1** $q_0 \neq 0$: If $q_0 \neq 0$ we get that $r_0 = r_1 = 0$ since otherwise the minimal degree of $qr < 2$, moreover $r_2 \neq 0$. This implies that $q_3 = 0, q_1 \neq 0$ (otherwise $\deg qr > 4$, and there would not be a homogeneous part of degree 3). It is left to check whether q_2 vanishes or not.

   (a) $q_2 \neq 0$: If $q_2 \neq 0$ then $r_3 = 0$ and we get that $r = r_2 = p_2$ is homogeneous of degree 2, and that p_2 divides p, in particular $(X_3 + X_4)$ divides $p(X_3, X_4)$. Considering this composition in $\overline{k}(X_3)[X_4]$ this implies that $-X_4$ is a root of $p(X_3, X_4)$ that is $p(X_3, -X_3) = 0$. The coefficient of the fourth degree of $p(X_3, -X_3)$ is then

$$(B^2 - 2a_3(-2a_2 + 2a_5 + B - C) - 2a_4(-2a_2 + 2a_5 + B - C) - C^2 + D)$$

   that vanishes for at most 2 values of C.

   (b) $q_2 = 0$: If $q_2 = 0$, then $r_3 \neq 0$ and we have that $q = q_0 + q_1, r = r_2 + r_3$, such that $q_0 r_2 = p_2, q_1 r_3 = p_4$ and $q_1 r_2 + q_0 r_3 = p_3$. Without loss of generality we may assume

$$q_1 = X_4, \ ((B + C - 2a_2)X_3 + 2a_4 X_4), \text{ or } \ (a_3 X_3 + a_4 X_4)$$
$$r_3 = \frac{p_4}{q_1}, \quad r_2 = \frac{p_2}{q_0}$$

   If $q_1 \neq X_4$ we get that X_4 divides r_3, p_3 and therefore $X_4 | r_2$ which is a contradiction. Therefore we get that $q_1 = X_4$, and

$$p_2(X_3, X_4)/q_0 + q_0((B + C - 2a_2)X_3 + 2a_4 X_4)(a_3 X_3 + a_4 X_4) =$$
$$2(B - C - 2a_2 + 2a_5)(a_4 X_4^2 + a_3 X_3^2) + (B^2 - C^2 + D)X_3 X_4$$

   comparing coefficients gives 3 equations for C, q_0 that have at most 6 solutions for C.

(2) **Case 2** $q_0 = 0$: If $q_0 = 0$ then $q_1 \neq 0$ (otherwise $\deg(r) \deg(q)$ that contradicts our assumption). This implies that $r_0 = 0, r_1, q_2 \neq 0$. If $q_3 \neq 0$ then $r_2 = 0$ and hence we get that r_1 divides p, p_2, p_3, p_4 which we saw above that can happen for at most 2 values of C. We therefore get that $q = q_1 + q_2, r = r_1 + r_2$ satisfying $q_1 r_1 = p_2, q_1 r_2 + q_2 r_1 = p_3, q_2 r_2 = p_4$. Since X_4 does not divide p_2 and does divide p_3, p_4 we find that X_4 must divide q_2, r_2. We therefore assume without loss of generality, that $q_2 = X_4(a_3 X_3 + a_4 X_4), r_2 = p_4/q_2$, and $(q_1, r_1) = (\mu(X_3 + X_4), \frac{1}{\mu}(2a_5(B - C + 2a_2)X_4 + 2a_3 X_3))$ or $(\mu(2a_5(B - C + 2a_2)X_4 + 2a_3 X_3), \frac{1}{\mu}(X_3 + X_4))$. Comparing coefficients again for $q_1 r_2 + q_2 r_1 = p_3$ we get 3 equations for C, μ that have at most 6 solutions in C.

Combining all restrictions for C we get that if C is outside a set of cardinality at most 14 $p(X_3, X_4)$ is irreducible. $\qquad\square$

REFERENCES

[1] E. Bombieri. *On exponential sums in nite elds. II.*, Invent. Math., 47(1):29-39, 1978.

[2] J. Bourgain, *A remark on Quantum Ergodicity for cat maps*, in: Geometric Aspects of Functional Analysis, Papers from the Israel Seminar (GAFA) held 2004-2005, Edited by V. D. Milman and G. Schechtman. Lecture Notes in Mathematics , Vol. 1910, Springer-Verlag, Berlin, 2007.

[3] Y. Colin de Verdiére. *Ergodicité et fonctions propres du laplacien* Comm. Math. Phys., 102(3):497-502, 1985.

[4] Mirko Degli Esposti, Sandro Graffi, and Stefano Isola. *Classical limit of the quantized hyperbolic toral automorphisms.* Comm. Math. Phys., 167(3):471-507, 1995.

[5] Pierre Deligne. *La conjecture de Weil. II.* Inst. Hautes Études Sci. Publ. Math., (52):137-252, 1980.

[6] Bruno Eckhardt, Shmuel Fishman, Jonathan Keating, Oded Agam, Jorg Main, and Kirsten Muller. *Approach to ergodicity in quantum wave functions.* Phys. Rev. E, 52(6):5893-5903, Dec 1995.

[7] Faure, Frédéric and Nonnenmacher, Stéphane and De Bièvre, Stephan. *Scarred eigenstates for quantum cat maps of minimal periods.* Comm. Math. Phys., 239(3):449-492, 2003.

[8] Mario Feingold and Asher Peres. *Distribution of matrix elements of chaotic systems.* Phys. Rev. A (3), 34(1):591-595, 1986.

[9] Shamgar Gurevich and Ronny Hadani. *Proof of the Kurlberg-Rudnick rate conjecture.* C. R. Math. Acad. Sci. Paris, 342(1):69-72, 2006.

[10] J. H. Hannay and M. V. Berry. *Quantization of linear maps on a torus-Fresnel diffraction by a periodic grating.* Phys. D, 1(3):267-290, 1980.

[11] R. Holowinsky and K. Soundararajan. *Mass equidistribution for hecke eigenforms.* preprint:arxiv.org/abs/0809.1636, 2008.

[12] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory, volume 53 of American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, RI, 2004.

[13] Nicholas M. Katz. *Sums of betti numbers in arbitrary characteristic.* Finite Fields and Their Applications, 7(1):29-44, 2001.

[14] J. P. Keating. *The cat maps: quantum mechanics and classical motion.* Nonlinearity, 4(2):309-341, 1991.

[15] Dubi Kelmer. *Arithmetic quantum unique ergodicity for symplectic linear maps of the multidimensional torus.* to appear in Annals of mathematics. preprint:arxiv.org/abs/math-ph/0510079.

[16] Dubi Kelmer. *Scarring on invariant manifolds for perturbed quantized hyperbolic toral automorphisms.* Comm. Math. Phys., 276(2):381-395, 2007.

[17] Pär Kurlberg. *On the order of unimodular matrices modulo integers.* Acta Arith., 110(2):141-151, 2003.

[18] Pär Kurlberg, Lior Rosenzweig, and Zeév Rudnick. *Matrix elements for the quantum cat map: Fluctuations in short windows.* Nonlinearity, 20(10):2289-2304, 2007.

[19] Pär Kurlberg and Zeév Rudnick. *Hecke theory and equidistribution for the quantization of linear maps of the torus.* Duke Math. J., 103(1):47-77, 2000.

[20] Pär Kurlberg and Zeév Rudnick. *On quantum ergodicity for linear maps of the torus.* Comm. Math. Phys., 222(1):201-227, 2001.

[21] Pär Kurlberg and Zeév Rudnick. *Value distribution for eigenfunctions of desymmetrized quantum maps.* Internat. Math. Res. Notices, (18):985-1002, 2001.

[22] Pär Kurlberg and Zeév Rudnick. *On the distribution of matrix elements for the quantum cat map.* Ann. of Math. (2), 161(1):489-507, 2005.

[23] Serge Lang and André Weil. *Number of points of varieties in Finite Fields.* Amer. J. Math., 76:819-827, 1954.

[24] Elon Lindenstrauss. *Invariant measures and arithmetic quantum unique ergodicity.* Ann. of Math. (2), 163(1):165-219, 2006.

[25] Wenzhi Luo and Peter Sarnak. *Quantum variance for Hecke eigenforms.* Ann. Sci. École Norm. Sup. (4), 37(5):769-799, 2004.

[26] Mœglin, Colette and Vignéras, Marie-France and Waldspurger, Jean-Loup *Correspondances de Howe sur un corps p-adique*, volume 1291 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1987.

[27] Lior Rosenzweig. *Fluctuations of the matrix elements of the quantum cat map*, Phd. thesis. In preperation.

[28] Zeév Rudnick. *The arithmetic theory of quantum maps. In Equidistribution in number theory, an introduction*, volume 237 of NATO Sci. Ser. II Math. Phys. Chem., pages 331-342. Springer, Dordrecht, 2007.

[29] A. I. Šnirel'man, *Ergodic properties of eigenfunctions.* Uspehi Mat. Nauk, 29(6(180)):181-182, 1974.

[30] K Soundararajan. *Quantum unique ergodicity for $SL_2(\mathbb{Z})\backslash H$.* preprint:arxiv.org/abs/0901.4060, 2009.

[31] André Weil. *On some exponential sums.* Proc. Nat. Acad. Sci. U.S.A., 34:204-207, 1948.

[32] Steven Zelditch. *Uniform distribution of eigenfunctions on compact hyperbolic surfaces.* Duke Math. J., 55(4):919-941, 1987.

Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

E-mail address: rosenzwe@post.tau.ac.il